Durham University

Durham Research Online

Deposited in DRO:

05 October 2018

Version of attached file:

Accepted Version

Peer-review status of attached file:

Peer-reviewed

Citation for published item:

Koulieris, George-Alex and Aksit, Kaan and Richardt, Christian and Mantiuk, Rafal (2018) 'Cutting-edge VR/AR display technologies (gaze-, accommodation-, motion-aware and HDR-enabled).', in SIGGRAPH Asia 2018 Courses. New York: Association for Computing Machinery (ACM), p. 5.

Further information on publisher's website:

https://doi.org/10.1145/3277644.3277771

Publisher's copyright statement:

© 2018 Copyright held by the owner/author(s) 2018. This is the author's version of the work. It is posted here for your personal use. Not for redistribution. The definitive Version of Record was published in SIGGRAPH Asia 2018 Courses https://doi.org/10.1145/3277644.3277771

Additional information:

Use policy

The full-text may be used and/or reproduced, and given to third parties in any format or medium, without prior permission or charge, for personal research or study, educational, or not-for-profit purposes provided that:

- a full bibliographic reference is made to the original source
- a link is made to the metadata record in DRO
- the full-text is not changed in any way

The full-text must not be sold in any format or medium without the formal permission of the copyright holders.

Please consult the full DRO policy for further details.

Durham University Library, Stockton Road, Durham DH1 3LY, United Kingdom
Tel : +44 (0)191 334 3042 | Fax : +44 (0)191 334 2971
https://dro.dur.ac.uk

Cutting-Edge VR/AR Display Technologies (Gaze-, Accommodation-, Motion-aware and HDR-enabled)

GEORGE-ALEX KOULIERIS, Durham University, United Kingdom
KAAN AKŞIT, NVIDIA, United States of America
CHRISTIAN RICHARDT, University of Bath, United Kingdom
RAFAŁ MANTIUK, University of Cambridge, United Kingdom

Near-eye (VR/AR) displays suffer from technical, interaction as well as visual quality issues which hinder their commercial potential. This tutorial will deliver an overview of cutting-edge VR/AR display technologies, focusing on technical, interaction and perceptual issues which, if solved, will drive the next generation of display technologies. The most recent advancements in near-eye displays will be presented providing (i) correct accommodation cues, (ii) near-eye varifocal AR, (iii) high dynamic range rendition, (iv) gaze-aware capabilities, either predictive or based on eye-tracking as well as (v) motion-awareness. Future avenues for academic and industrial research related to the next generation of AR/VR display technologies will be analyzed.

Additional Information:
Course website: https://vrdisplays.github.io/sigasia2018/

Authors' addresses: George-Alex Koulieris, Durham University, United Kingdom, georgios.a.koulieris@durham.ac.uk; Kaan Akşit, NVIDIA, United States of America, kaksit@nvidia.com; Christian Richardt, University of Bath, United Kingdom, christian@richardt.name; Rafał Mantiuk, University of Cambridge, United Kingdom, rafal.mantiuk@cl.cam.ac.uk.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).
© 2018 Copyright held by the owner/author(s).
ACM 978-1-4503-6026-5/18/12
https://doi.org/10.1145/3277644.3277771

1 COURSE CONTENT

1.1 Accommodation-aware VR

Head-mounted Displays (HMDs) often provoke discomfort and nausea. Recent exciting work has showcased that when accommodation and vergence distances match in an HMD, comfort significantly improves [Koulieris et al. 2017]. One way to achieve such a match is by combining gaze-contingent, depth-of-field (DoF) rendering with new developments on focus-adjustable lenses or spatial light modulators (SLMs). In this tutorial, the latest advancements on adjustable lenses and SLMs will be examined that provide correct accommodation cues depending on the distance of the object being observed in the virtual scene.

1.2 Near-Eye Varifocal AR

New advancements in display engineering and a broader understanding of vision science have led to computational displays for VR and AR. Today, such displays promise a more realistic and comfortable experience through techniques such as lightfield displays, holographic displays, always-in-focus displays, multiplane displays, and varifocal displays. In this talk, new optical layouts for see-through computational near-eye displays are presented that are simple, compact, varifocal, and provide a wide field of view with clear peripheral vision and large eyebox [Akşit et al. 2017]. Key to research efforts so far contain novel see-through rear-projection holographic screens and deformable mirror membranes [Dunn et al. 2017]. Fundamental trade-offs are established between the quantitative parameters of resolution, field of view and the form-factor of the designs; opening an intriguing avenue for future work on accommodation-supporting AR displays.

1.3 HDR-enabled

Currently, commercial HMDs are based on standard dynamic range (SDR) imaging systems. High dynamic range display and rendering technologies, capable of depicting the extreme brightness range and an extensive range of colours, could improve visual quality, enhancing immersion and sense of realism [Mantiuk et al. 2015]. The course will analyze recent developments in relation to high dynamic range content production, rendering and display [Mantiuk et al. 2008] and how this can be incorporated in VR displays. It will analyze the challenges of introducing higher brightness levels to VR and the effect it could have on visual quality and comfort.

1.4 Motion-aware

Existing HMDs provide limited input to a user beyond the positional tracking of the HMD and/or controllers. Users currently cannot see or perceive their own body in VR [Rhodin et al. 2016b]. This course will present experiments conducted with a novel head-mounted marker-less motion capture system in immersive VR applications [Rhodin et al. 2016a]. The system comprises of two fish-eye cameras attached to an HMD, tracking the motion of a user wearing it. By utilizing such as lightweight capture rig, geared for HMD-based VR, egocentric motion capture is feasible. Applications will be demonstrated in which the user looks down at their virtual self. Current HMD-based systems only track the pose of the display. The tutorial will showcase novel approaches adding motion capture of the wearer's full body, evoking a higher level of immersion.

2 COURSE HISTORY AND RELEVANT EXISTING COURSES

This is a new course on a topic that has so far not been covered at SIGGRAPH. While significant advances in VR/AR display technologies have been made in the past five years, less has been specifically written about the state of the art in display technologies. We hope to address this with our course.

Some aspects related to our course topics have been covered in previous SIGGRAPH courses:

- Applications of visual perception to virtual reality rendering[1] by Anjul Patney, Marina Zannoli, George-Alex Koulieris, Joohwan Kim, Gordon Wetzstein and Frank Steinicke (SIGGRAPH 2017)
  — Considered the role of ongoing and future research in visual perception to improve rendering for virtual reality. While we will mention perceptual issues, they will not be the main focus of our course. Instead, we focus on display technologies themselves, particularly hardware architectures.
- Build your own VR system: an introduction to VR displays and cameras for hobbyists and educators[2] by Gordon Wetzstein, Robert Konrad, Nitish Padmanaban and Hayato Ikoma (SIGGRAPH 2017)
  – Introduces basic concepts regarding design and programming of existing VR/AR displays. The proposed course will go beyond the existing technologies and will focus on the technologies we will find in the VR/AR headsets in the near future.
- Augmented reality: principles and practice[3] by Dieter Schmalstieg and Tobias Höllerer (SIGGRAPH 2016)
  — The main focus of this course was augmented reality; we will be focusing on the optical design of both virtual and augmented reality devices.
- Put on your 3D glasses now: the past, present, and future of virtual and augmented reality[4] by Douglas Lanman, Henry Fuchs, Mark Mine, Ian McDowall, and Michael Abrash (SIGGRAPH 2014)
  — A comprehensive survey of VR only display technologies, with a strong focus on head-mounted displays. However, significant advances in optical design, hardware and interaction in VR/AR have occurred in the last 4 years which we hope to address in our course.

3 COURSE SCOPE

In our course we focus on the technical, interaction and perceptual issues of VR/AR display technologies that, if solved, will drive the next generation of display technologies. In particular we cover the most recent advancements in near-eye displays such as displays providing correct accommodation cues, high dynamic range rendition, gaze and motion awareness etc.

3.1 Intended audience

As VR/AR technologies are becoming ubiquitous, our course is targeted at a broad audience such as students, academics and professionals wishing to gain an understanding of how near-eye displays for VR/AR headsets work and benefit from the background to current state-of-the-art systems and the problems currently being tackled to bring VR/AR displays to wide use.

3.2 Prerequisites, Pedagogic Intentions and Methods

We expect both beginners and experienced people in the field will find the course engaging, as useful insights from visual perception and optical design for near-eye displays will be presented. A basic knowledge of computer graphics is useful. Schematic diagrams, photographs, animations and videos will be employed to facilitate explanation and learning. The syllabus is “not too easy, but not too difficult”. We hope to maintain attendees in a state of learning “flow” by varying the level of difficulty from easy to hard and back, keeping them in an optimal learning zone without getting them bored or disappointed.

[1] https://doi.org/10.1145/3084873.3086551
[2] https://doi.org/10.1145/3084873.3084928
[3] https://doi.org/10.1145/2897826.2927365
[4] https://doi.org/10.1145/2614028.2628332

4 COURSE PRESENTER INFORMATION

George-Alex Koulieris — *Durham University*
georgios.a.koulieris@durham.ac.uk • https://koulieris.com

George-Alex Koulieris (B.Sc. in Computer Science and Telecommunications, University of Athens, M.Sc. in Computer Science, University of Economics and Business, Athens, PhD in Electronic & Computer Engineering, Technical University of Crete, Greece) is an Assistant Professor in the Dept. of Computer Science at Durham University. Before that he was a post-doctoral researcher at Inria, France, team GraphDeco, working on near-eye, stereo displays. Previously, he was a visiting scholar at UC Berkeley, working on the vergence – accommodation conflict for head-mounted displays. During his PhD studies he worked on gaze prediction for game balancing, level-of-detail rendering and stereo grading. He has previously co-organized two SIGGRAPH courses (Attention-aware rendering, mobile graphics and games in 2014, Applications of visual perception to virtual reality rendering in 2017).

Kaan Akşit — *NVIDIA*
kaksit@nvidia.com • https://kaanaksit.com

Kaan Akşit(B.S. in Electrical Engineering, Istanbul Technical University, M.Sc. in Electrical Power Engineering, RWTH Aachen University, Germany, Ph.D. in Electrical Engineering, Koç University, Turkey). In 2009, he joined Philips Research at Eindhoven, the Netherlands as an intern. In 2013, he joined Disney Research, Zurich, Switzerland as an intern. His past research includes topics such as visible light communications, optical medical sensing, solar cars, and auto-stereoscopic displays. Since July 2014, he is working as a research scientist at Nvidia Corporation located at Santa Clara, USA, tackling the problems related to computational displays for virtual and augmented reality.

Christian Richardt — *University of Bath*
christian@richardt.name • https://richardt.name

Christian Richardt is a Lecturer (=assistant professor) at the University of Bath. He received a BA and PhD in Computer Science from the University of Cambridge in 2007 and 2012, respectively. He was previously a postdoctoral researcher at Inria Sophia Antipolis, Max Planck Institute for Informatics and the Intel Visual Computing Institute. His research combines insights from vision, graphics, and perception to extract and reconstruct visual information from images and videos, to create high-quality visual experiences with a focus on 360° video, light fields and user-centric applications. He has previously co-organized two SIGGRAPH courses (User-Centric Videography in 2015, Video for Virtual Reality in 2017).

Rafał K. Mantiuk — *University of Cambridge*
rafal.mantiuk@cl.cam.ac.uk • https://www.cl.cam.ac.uk/~rkm38/

Rafał Mantiuk (PhD in Computer Science, Max-Planck-Institute for Computer Science) is a senior lecturer at the Department of Computer Science and Technology (Computer Laboratory), University of Cambridge (UK). His recent interests focus on designing rendering and display algorithms that adapt to human visual performance and viewing conditions in order to deliver the best images given limited resources, such as computation time, bandwidth or dynamic range. He contributed to early work on high dynamic range imaging, including quality metrics (HDR-VDP), video compression and tone-mapping. In 2017 he was awarded an ERC Consolidator grant to work on perceptual encoding of high dynamic range light fields.

5 COURSE SCHEDULE

(1) **Welcome and Introduction**
George Alex Koulieris, Durham University, 10 minutes
- Motivation: understand current VR/AR display challenges
  Overview:
- Learn how challenges relate to visual perception
- What can we do about them?
- Discover the state-of-the-art in relevant research

(2) **Multifocal Displays**
George Alex Koulieris, Durham University, 40 minutes
- Basic optics, accommodation, VA conflict, discomfort, performance
- Multi-focal display technologies

(3) **Near-eye VR/AR Display Technologies**
Kaan Aksit, NVIDIA, 40 minutes
- Optics for AR
- Varifocal AR

(4) **Coffee break**
15 minutes

(5) **HDR, Displays & Low-level Vision**
Rafal Mantiuk, Cambridge University, 40 minutes
- Display technologies
- High Dynamic Range (HDR) Rendering
- HDR in VR

(6) **Motion-aware Displays**
Christian Richardt, University of Bath, 40 minutes
- Motion-aware displays
- Perception of immersion
- Tracking in VR and AR
- Hand input devices
- Motion capture

(7) **Coffee break**
10 minutes

(8) **Demos and Summary**
All, 30 minutes

= Total time of 3 hours and 45 minutes

REFERENCES

Kaan Akşit, Ward Lopes, Jonghyun Kim, Peter Shirley, and David Luebke. 2017. Near-eye varifocal augmented reality display using see-through screens. *ACM Transactions on Graphics (TOG)* 36, 6 (2017), 189.

David Dunn, Cary Tippets, Kent Torell, Henry Fuchs, Petr Kellnhofer, Karol Myszkowski, Piotr Didyk, Kaan Akşit, and David Luebke. 2017. Membrane AR: varifocal, wide field of view augmented reality display from deformable membranes. In *ACM SIGGRAPH 2017 Emerging Technologies*. ACM, 15.

George-Alex Koulieris, Bee Bui, Martin Banks, and George Drettakis. 2017. Accommodation and Comfort in Head-Mounted Displays. *ACM Transactions on Graphics* 36, 4 (2017), 11.

Rafał Mantiuk, Scott Daly, and Louis Kerofsky. 2008. Display adaptive tone mapping. In *ACM Transactions on Graphics (TOG)*, Vol. 27. ACM, 68.

Rafał K Mantiuk, Karol Myszkowski, and Hans-Peter Seidel. 2015. *High dynamic range imaging*. Wiley Online Library.

Helge Rhodin, Christian Richardt, Dan Casas, Eldar Insafutdinov, Mohammad Shafiei, Hans-Peter Seidel, Bernt Schiele, and Christian Theobalt. 2016a. EgoCap: egocentric marker-less motion capture with two fisheye cameras. *ACM Transactions on Graphics (TOG)* 35, 6 (2016), 162.

Helge Rhodin, Nadia Robertini, Dan Casas, Christian Richardt, Hans-Peter Seidel, and Christian Theobalt. 2016b. General automatic human shape and motion capture using volumetric contour cues. In *European Conference on Computer Vision*. Springer, 509–526.

Received July 2018

SIGGRAPH ASIA 2018 TOKYO

Cutting-Edge VR/AR Display Technologies

(Gaze-, Accommodation-, Motion-aware and HDR-enabled)

George-Alex Koulieris

Kaan Akşit

Christian Richardt

Rafał Mantiuk

Durham University

NVIDIA

University of Bath

University of Cambridge

Course at a glance

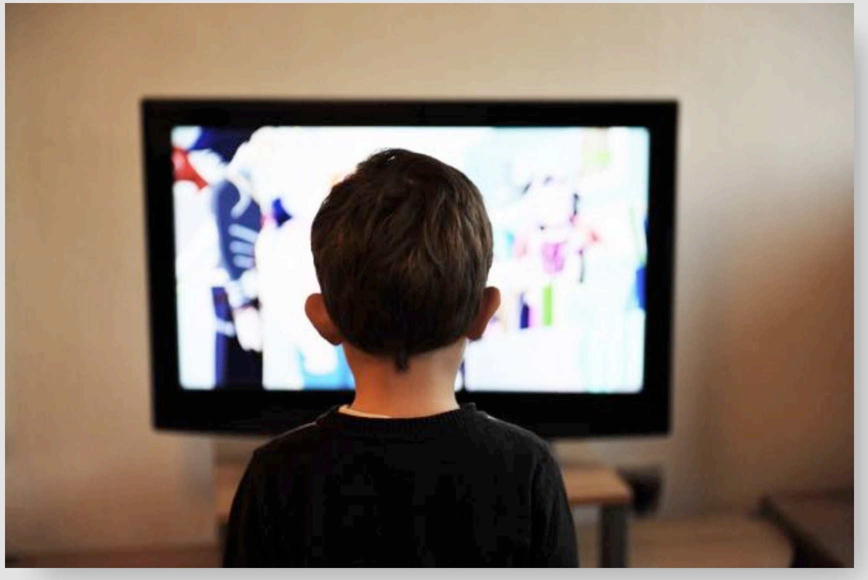
- Understand current VR/AR display challenges
- Learn how challenges relate to visual perception
- What can we do about them?
- Discover the state-of-the-art in relevant research

Speakers

- Kaan Akşit, NVIDIA, USA
- Christian Richardt, University of Bath, UK
- Rafał Mantiuk, University of Cambridge, UK
- George-Alex Koulieris, Durham University, UK

Let's get started

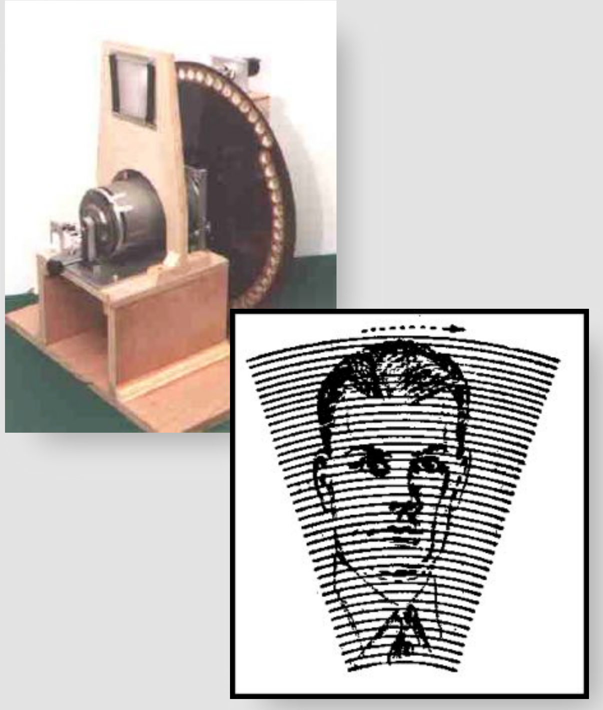
A Turing test for displays



George-Alex Koulieris

Displays

- Displays are virtual windows to remote scenes
- We have gone far from the Nipkow disk …



George-Alex Koulieris

Virtual, augmented, mixed reality displays

George-Alex Koulieris

- Collectively, near-eye displays
- Immersion into virtual/augmented world
- Response to head motion
- Allows object manipulation/interaction

VR/AR/MR applications

- **Education**
- **Communication**
- **Healthcare**
- **Entertainment**
- **Manufacturing**
- Aviation
- Business
- Design
- Gaming
- Marketing
- Shopping
- Sports
- Travel
- Therapy

Near-eye displays market explosion

- Top companies involved
- Market flooded with devices
- *"A billion people in virtual reality"*
  Mark Zuckerberg, 2017
- Research surge:
  SIGGRAPH, IEEE VR, ISMAR, ...

George-Alex Koulieris

Before this becomes commonplace…

George-Alex Koulieris

Current display challenges

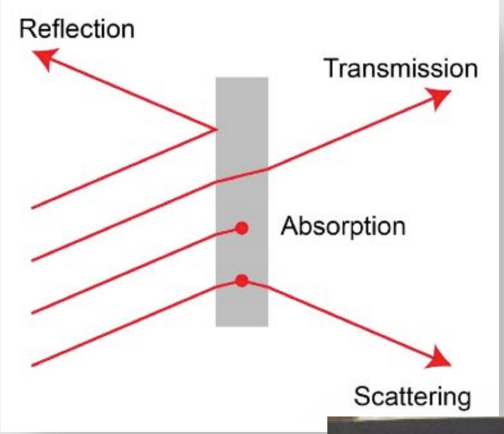
- Ergonomic / Comfort
- Visual Quality issues
- Perceptual
- Technical
- Interaction

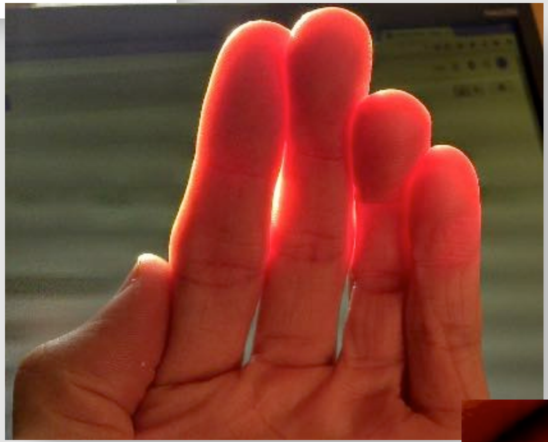
Exploiting knowledge from visual perception

- Display hardware and algorithms limited
- Produce different to natural light patterns
- Luckily, human visual system (HVS) limited
- Requirements restricted by HVS capabilities
- Visual perception as the optimizing function
- Achieve *perceptual effectiveness*
- Avoiding under-/over-engineering displays

But how do we take knowledge from visual perception into account?

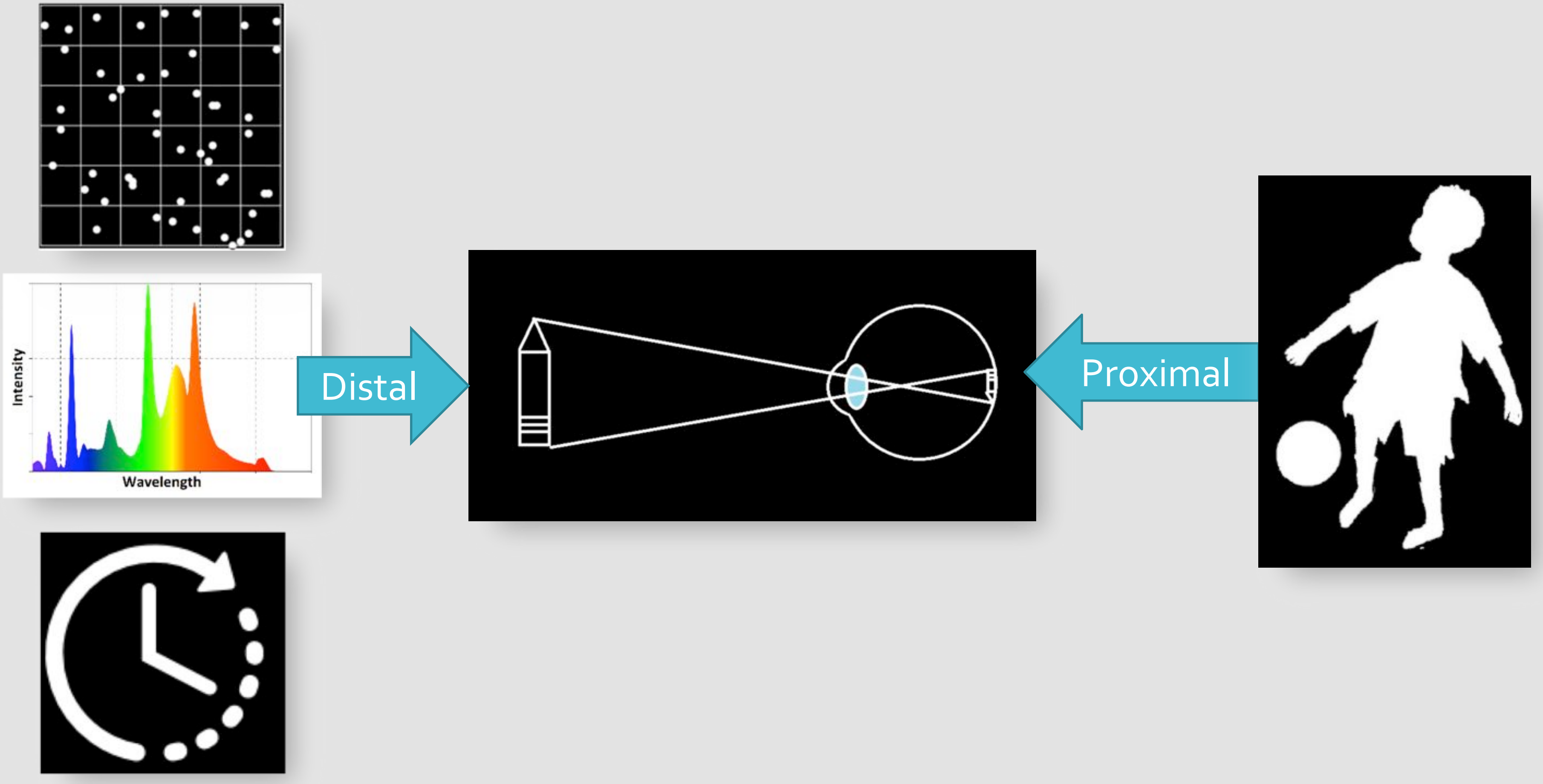
Human vision





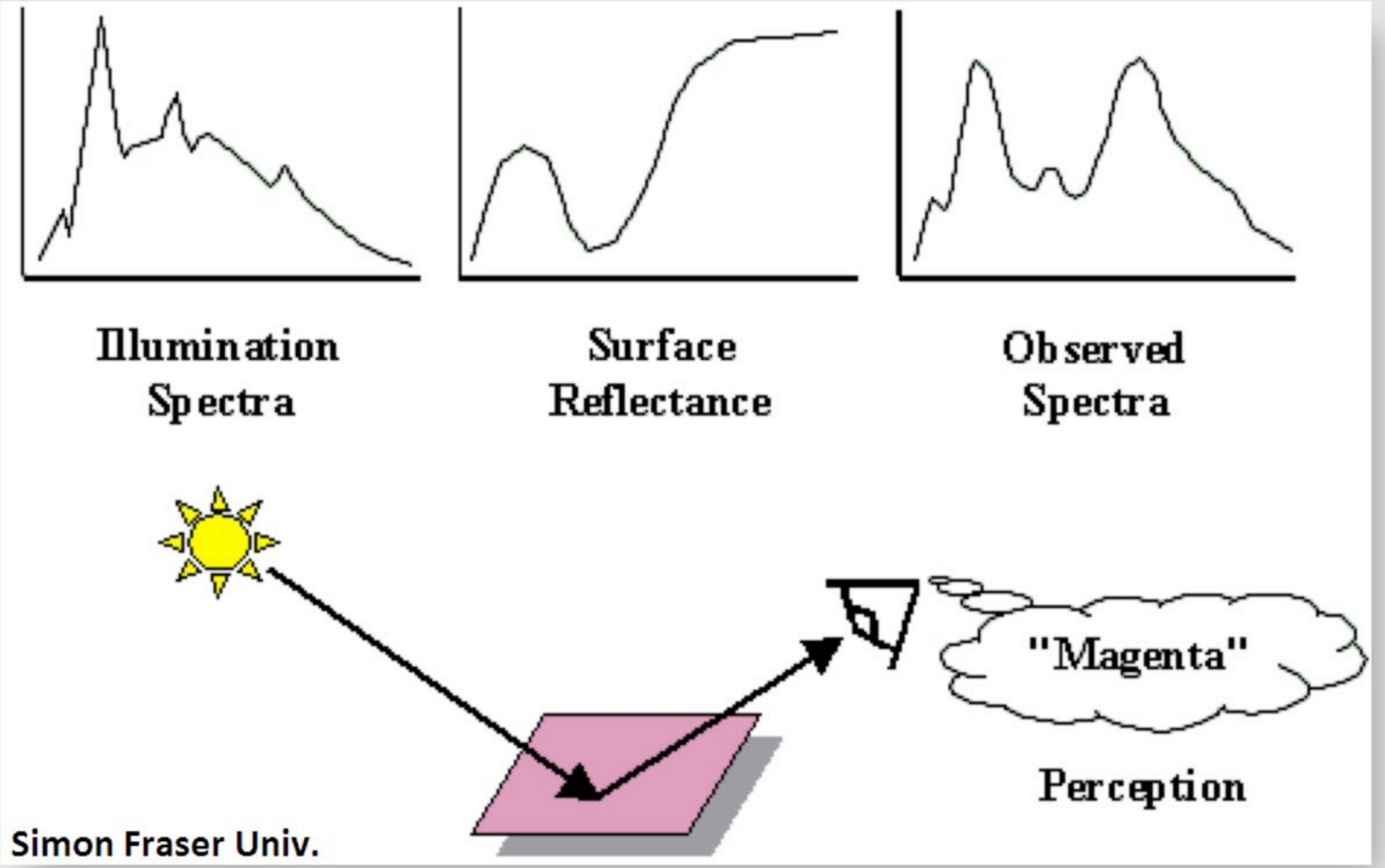
George-Alex Koulieris

What visual perception does

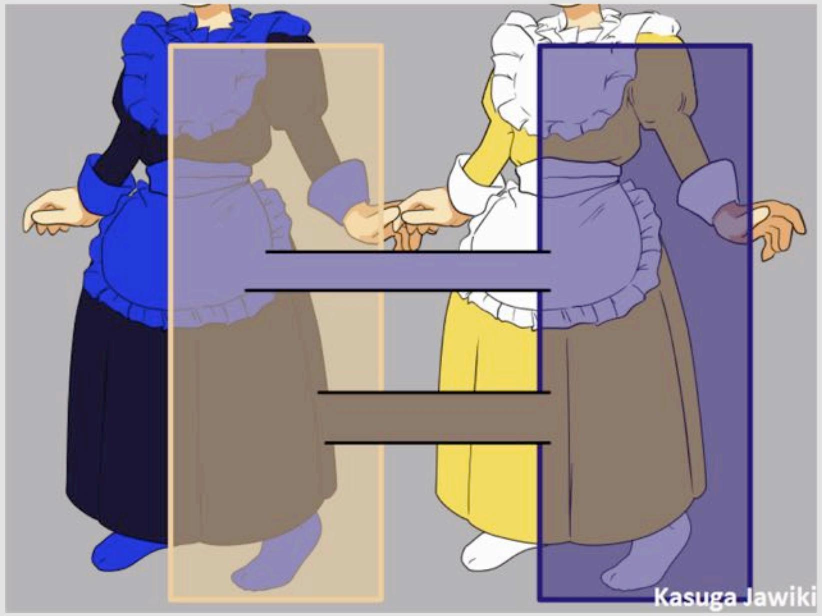


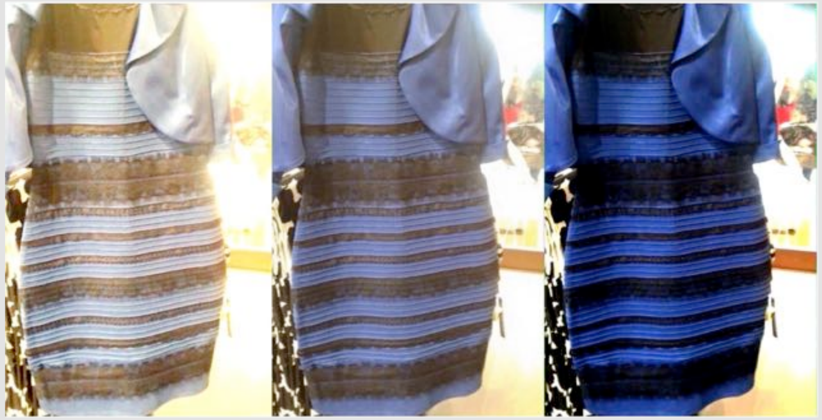
George-Alex Koulieris

Proximal → Distal: A difficult, inverse problem



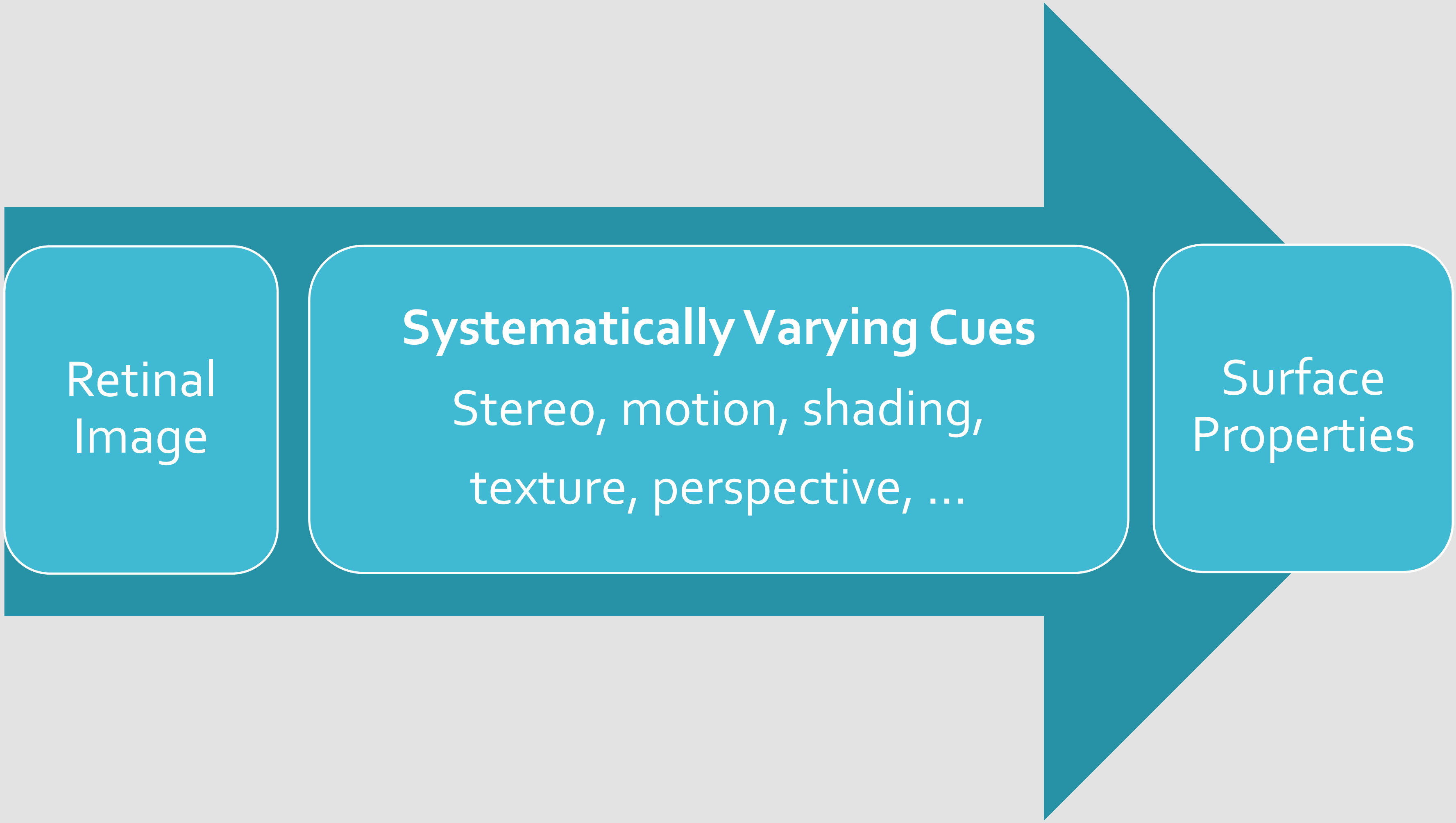
#thedress





George-Alex Koulieris

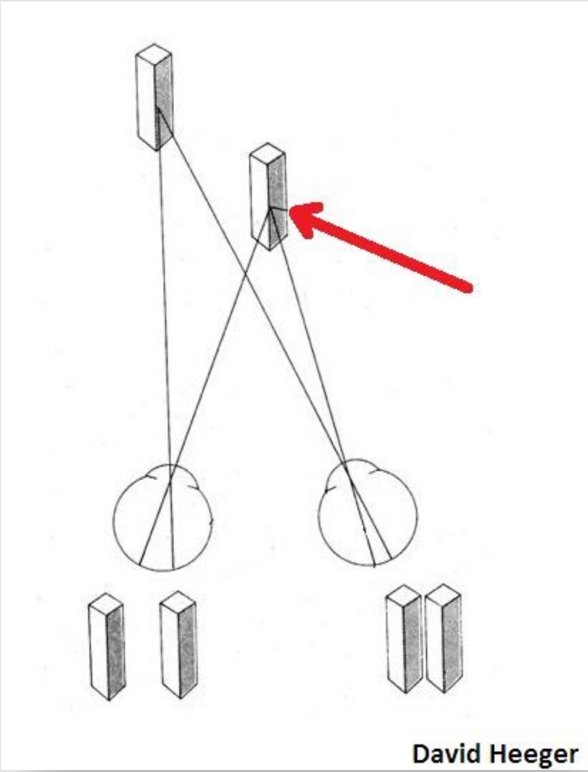
Visual perception and visual cues

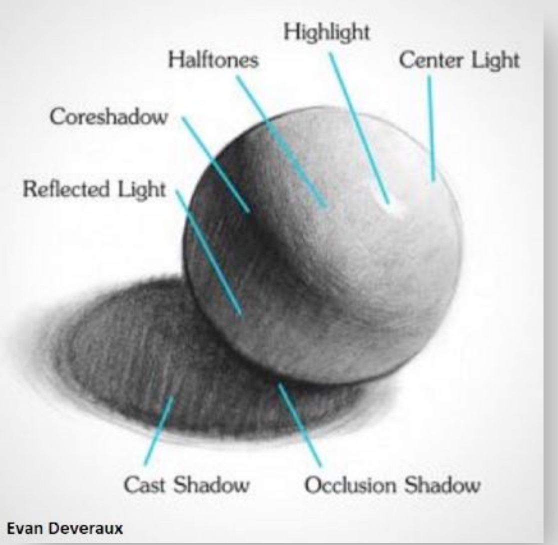


George-Alex Koulieris

Examples of visual cues

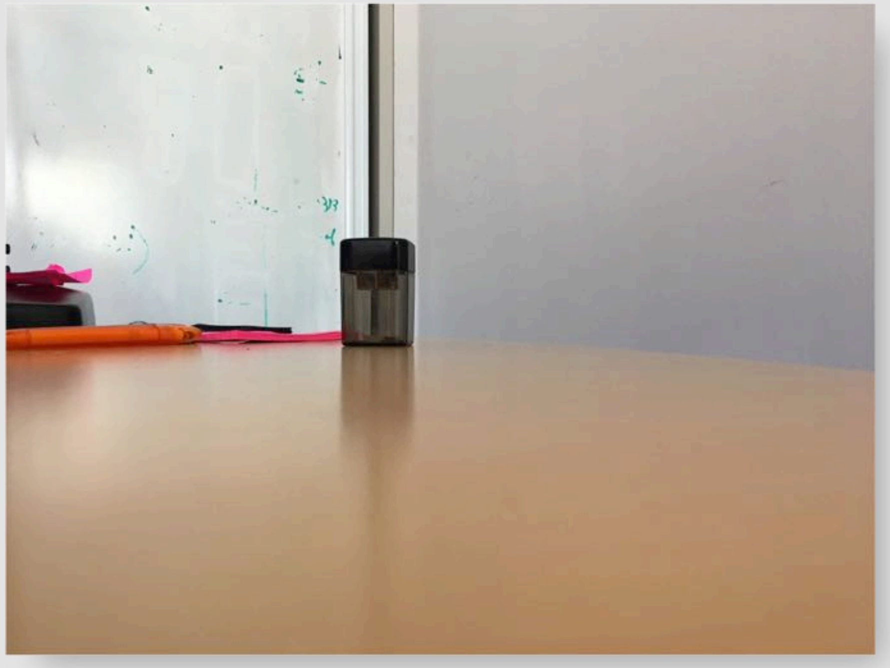
- Ocular-motor cues
  - eye position, focus
- Binocular disparity cues
- Motion cues
  - world, viewer
- Pictorial cues (monocular)
  - familiar size
  - relative size
  - shading
  - texture gradients
  - occlusion
  - …





George-Alex Koulieris

Cue integration

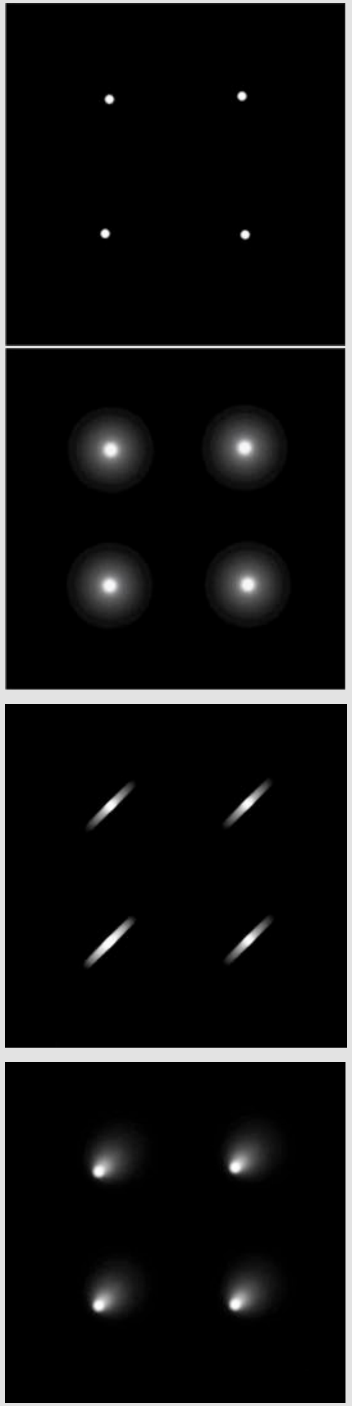


- Cues expected to **co-vary** for same environmental properties
- Expected consistent information overlap

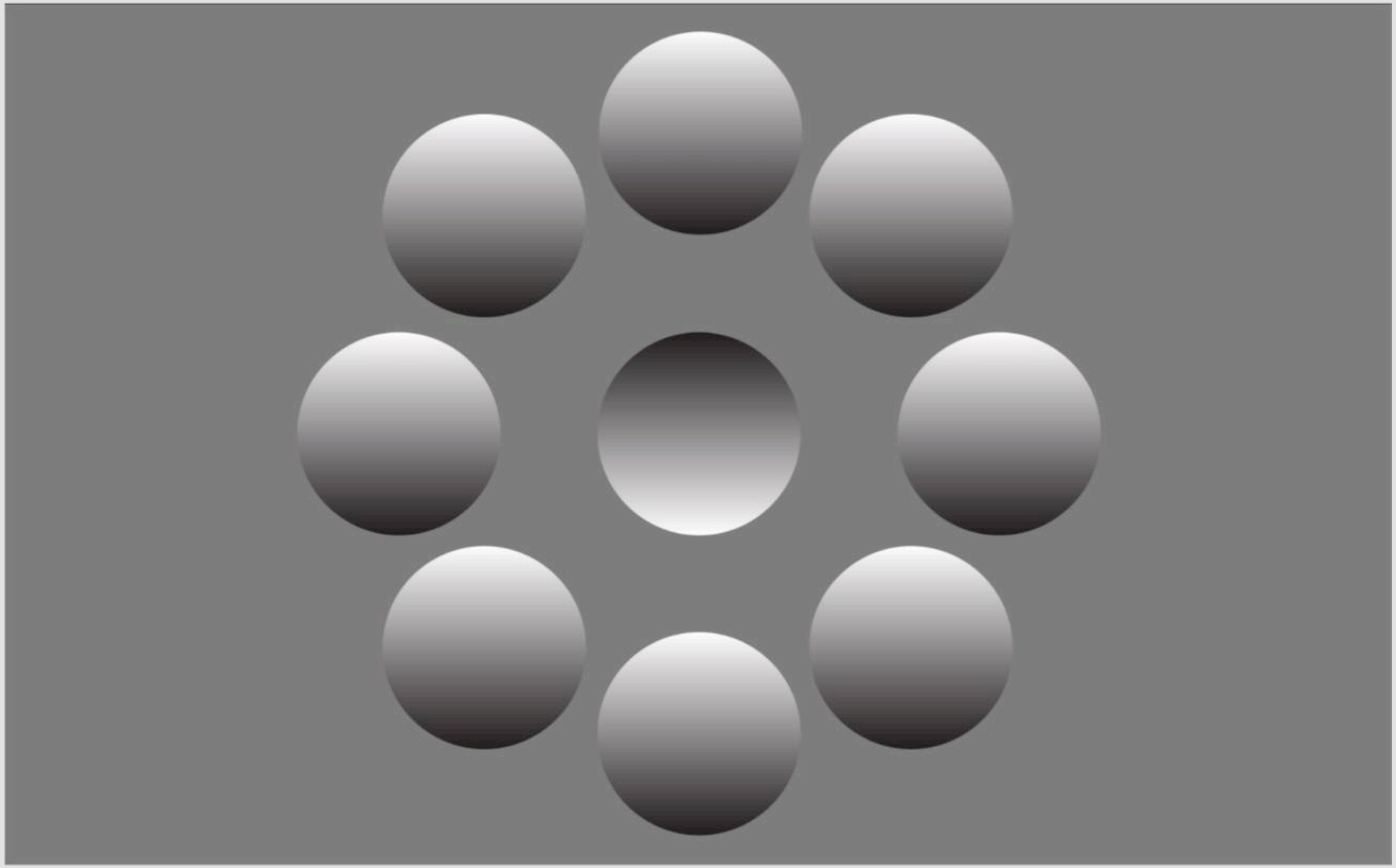
George-Alex Koulieris

Cue conflicts

- Cues often conflicting due to
  - VS errors
  - incomplete information (e.g., displays)
  - incorrect assumptions about the natural environment



Fun fact: conflicting cues



George-Alex Koulieris

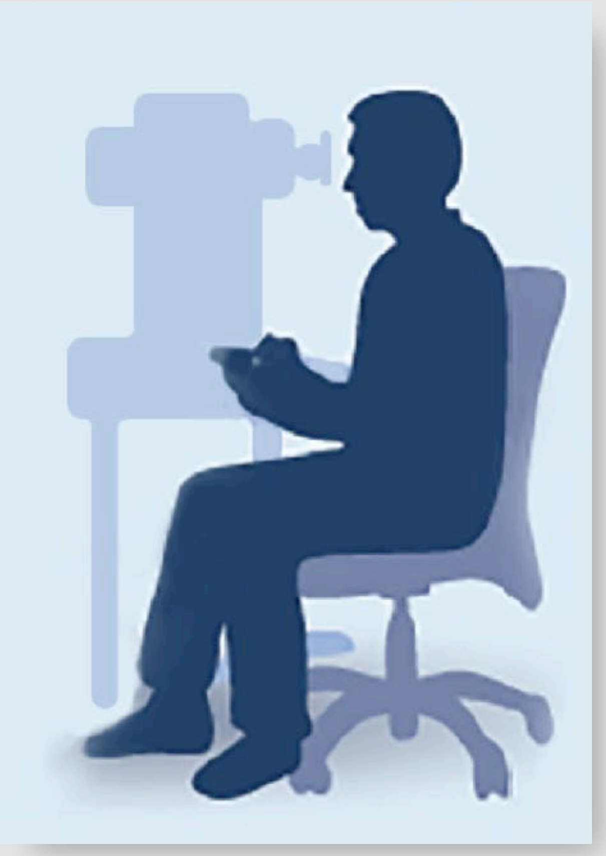
How do we study visual perception?

Psychophysical methods of study (1)

1. Show visual stimuli
2. Ask simple questions
3. Vary stimuli
4. GOTO 1

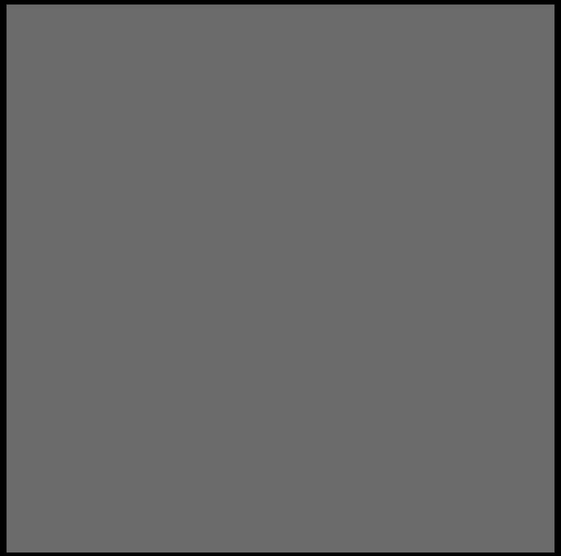
Psychophysical methods of study (2)

- N-A Forced choice tasks
- Method of adjustment
- Ascending/descending limits
- Staircase
- Constant stimuli



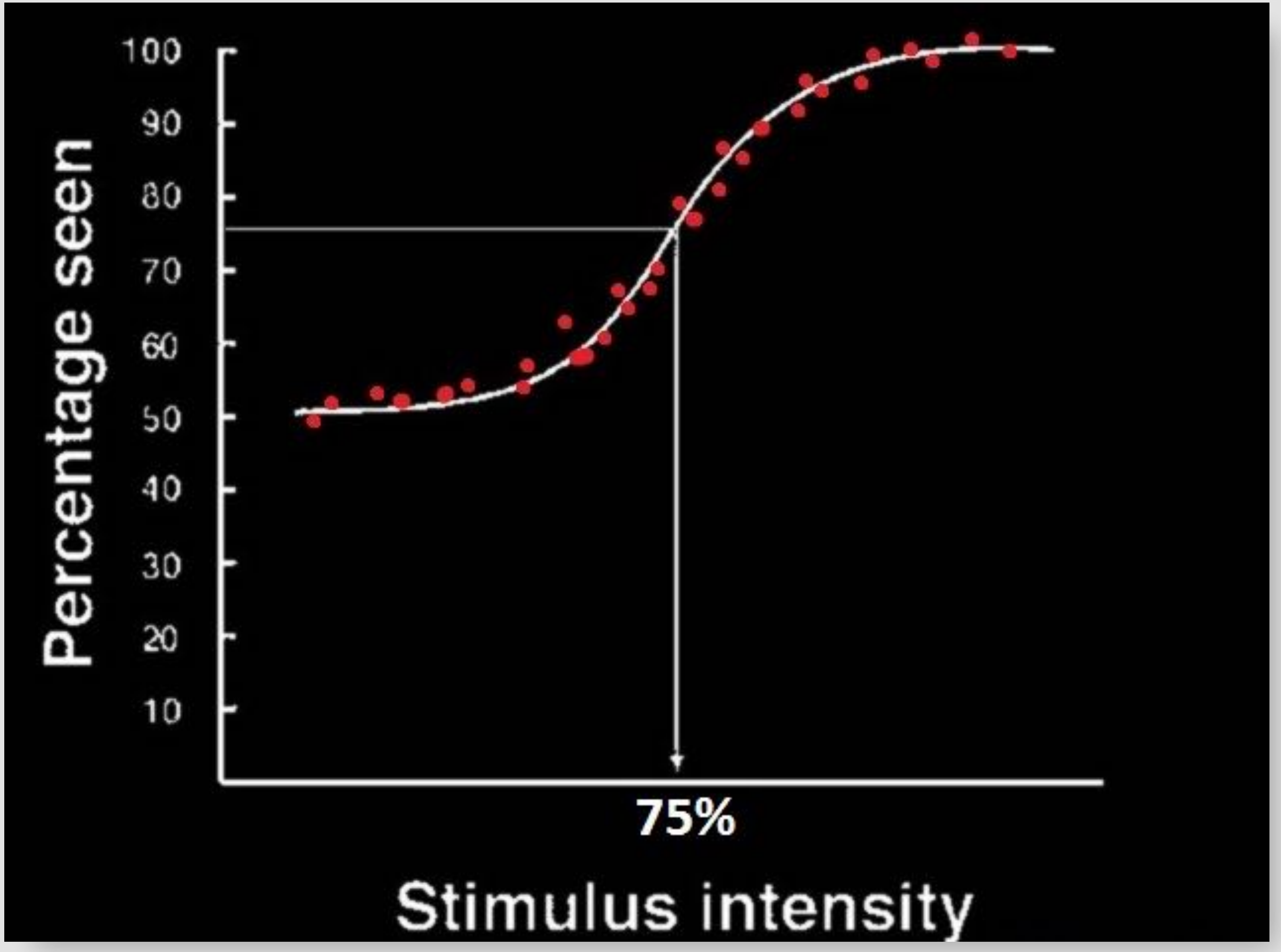
George-Alex Koulieris

Example: luminance threshold detection



- 200ms patch
- PRESENT / ABSENT ?

Psychometric functions



75% is half-way between chance and perfect performance!

Course take-aways

Course take-aways (1)

Q: Why multifocal displays?

Q: Why varifocal AR?

A: Eyes evolved to focus on objects.

Kaan Akşit

NVIDIA

George Alex Koulieris

Durham University

Course take-aways (2)

Q: Why HDR-enabled displays?

A: Relates to the sensitivity of the eyes.

Rafał Mantiuk

UNIVERSITY OF CAMBRIDGE

George-Alex Koulieris

Course take-aways (3)

Q: Why motion-aware displays?

A: Eyes attached on moving bodies.

Christian Richardt

UNIVERSITY OF BATH

George-Alex Koulieris

Summary

- Near-eye displays are beneficial to society
- Addressing challenges yields tremendous gains
- Near-eye displays a hot area for years to come
- Improving quality of experience in near-eye displays is an inter-disciplinary effort

- Questions so far?

SIGGRAPH ASIA 2018 TOKYO

George Alex Koulieris

Multi-focal Displays

SIGGRAPH Asia Course on Cutting-Edge VR/AR Display Technologies

Durham University

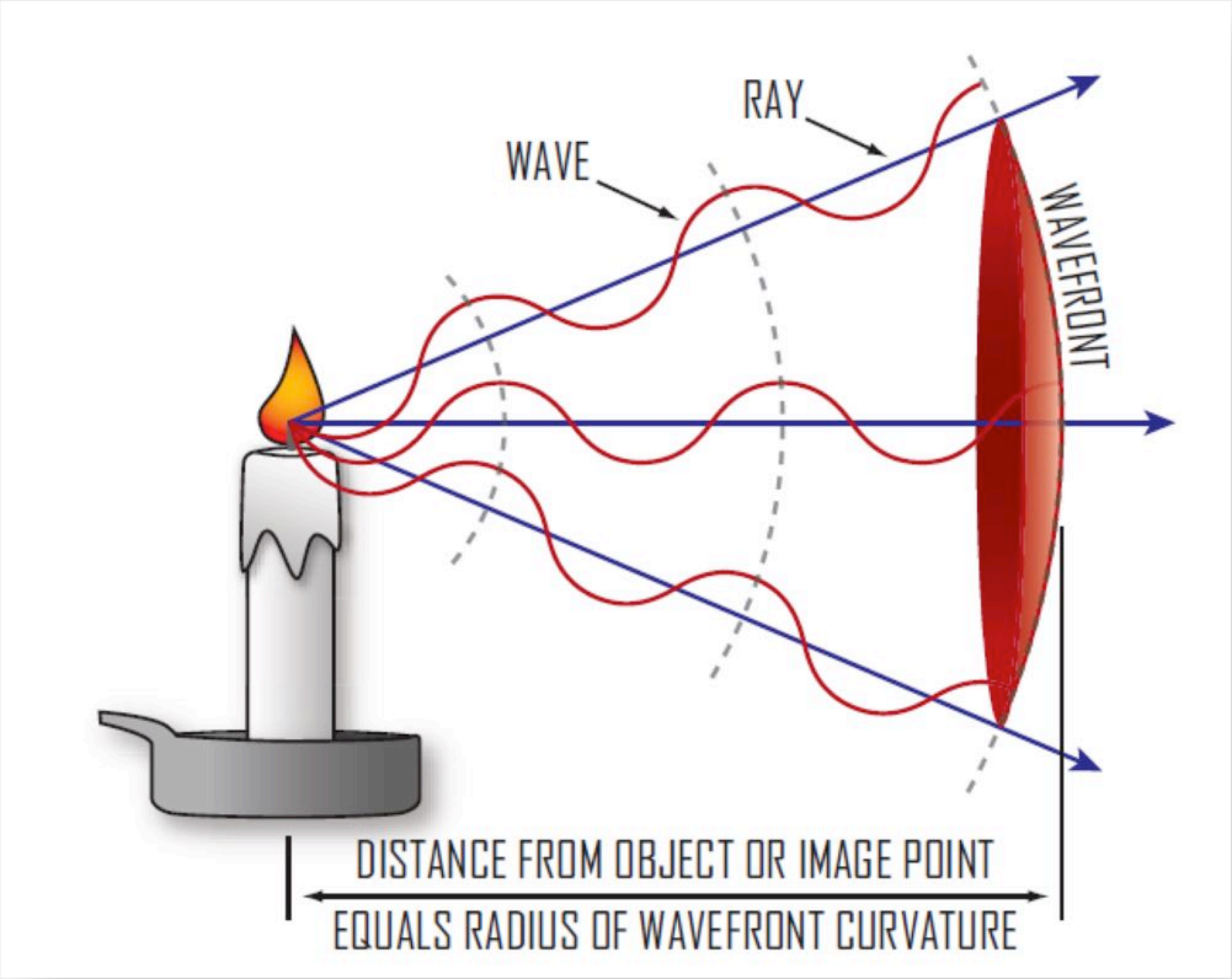
koulieris.com
@GeorgeKoulieris

At a glance

- Part 1: basic optics, accommodation, VA conflict, discomfort, performance
- Part 2: multi-focal display technologies

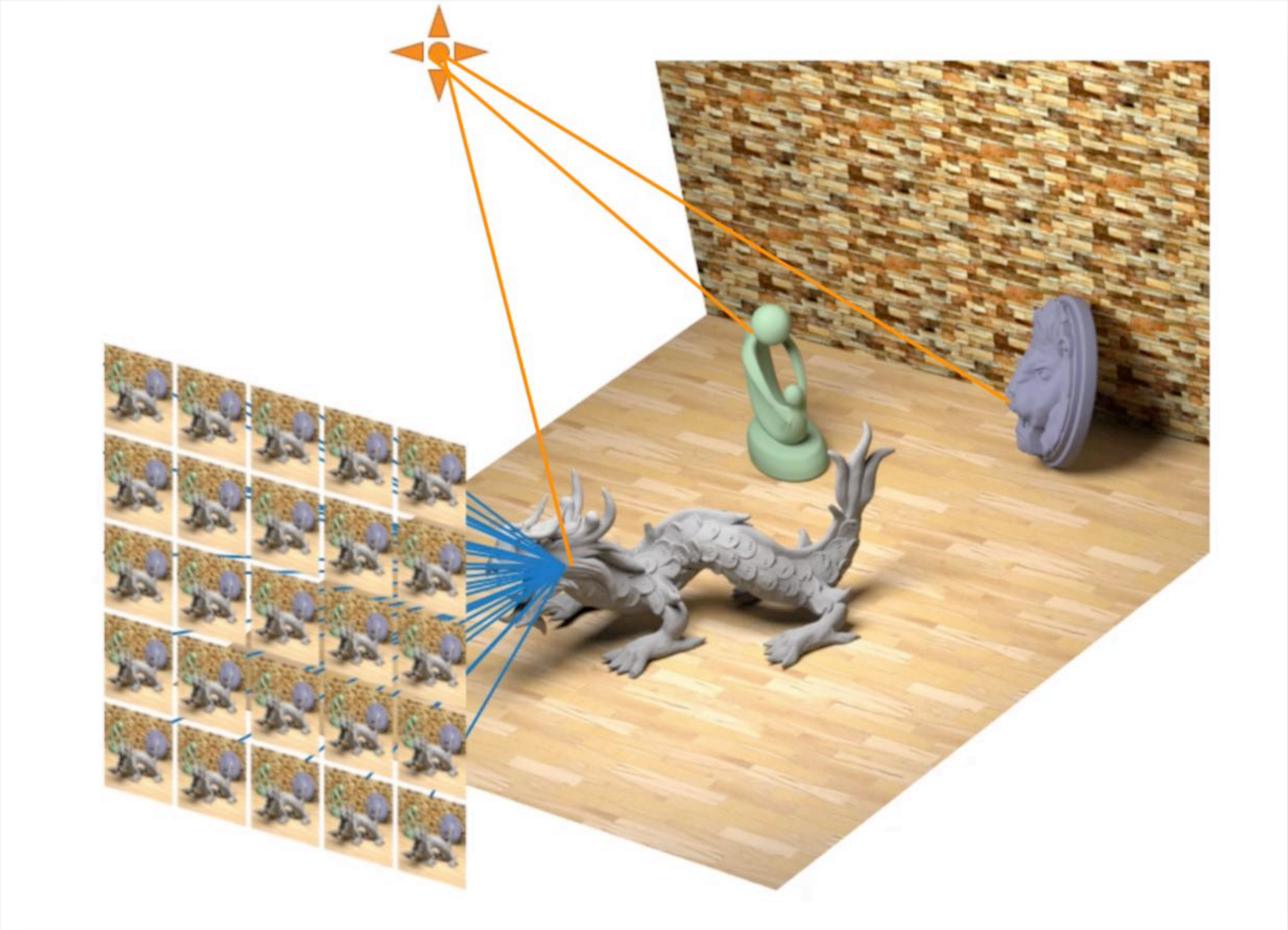
Part 1:
The basics

Light wave-front



Charle Laas

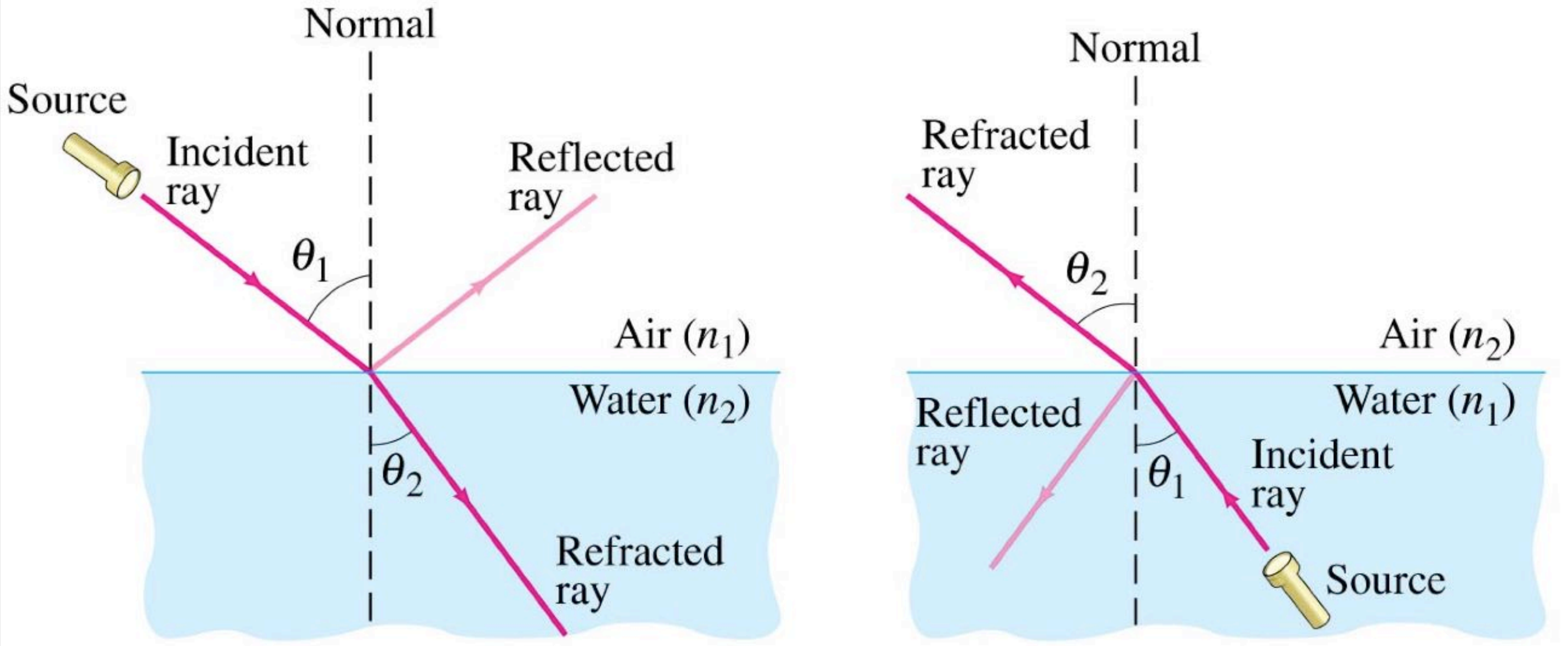
Natural light fields



Adapted from Mihara, 2016

George-Alex Koulieris

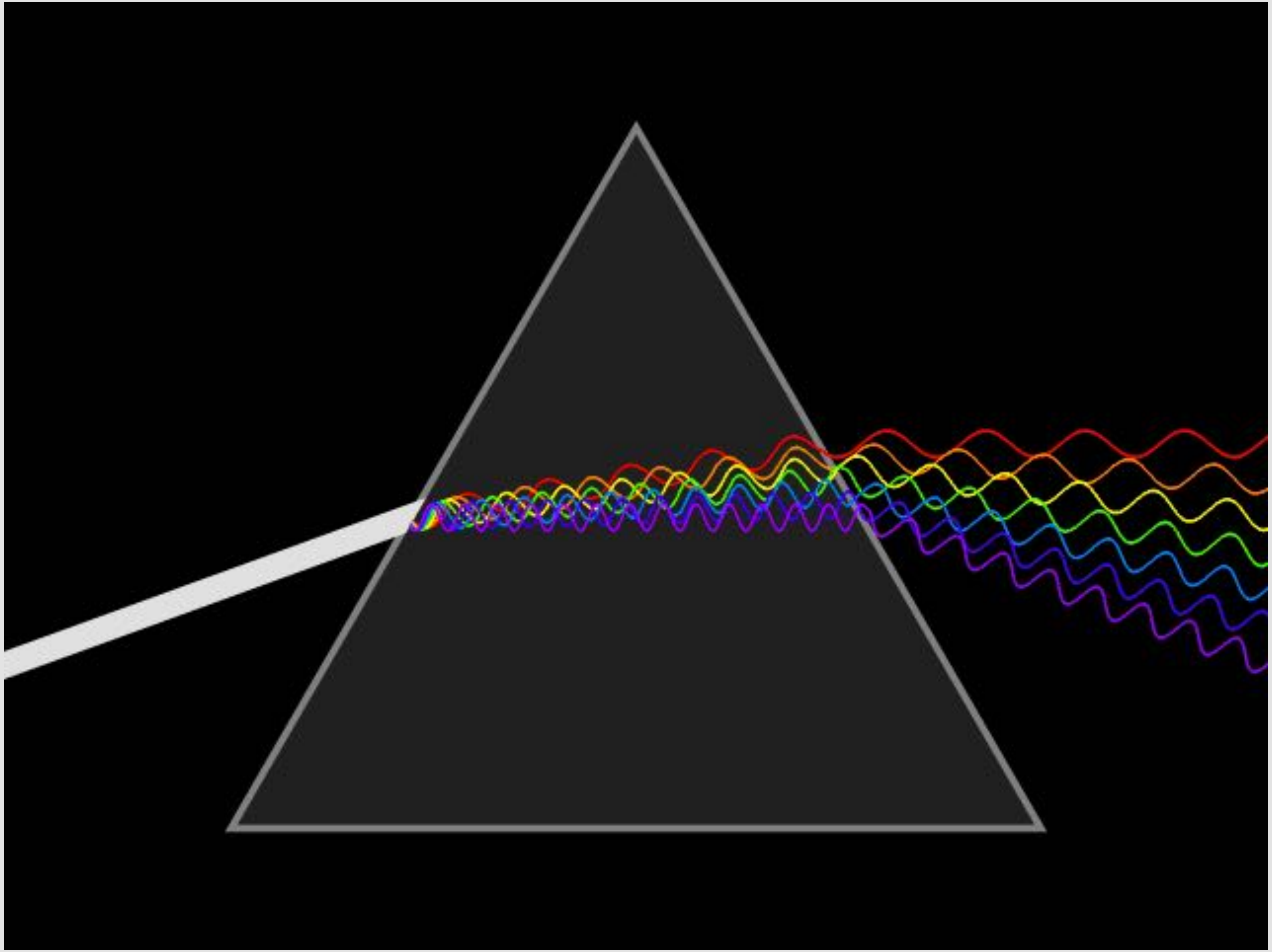
Refraction: Snell's law



(a) $n_2 > n_1$: Ray bends toward $\perp$ (b) $n_1 > n_2$: Ray bends away from $\perp$

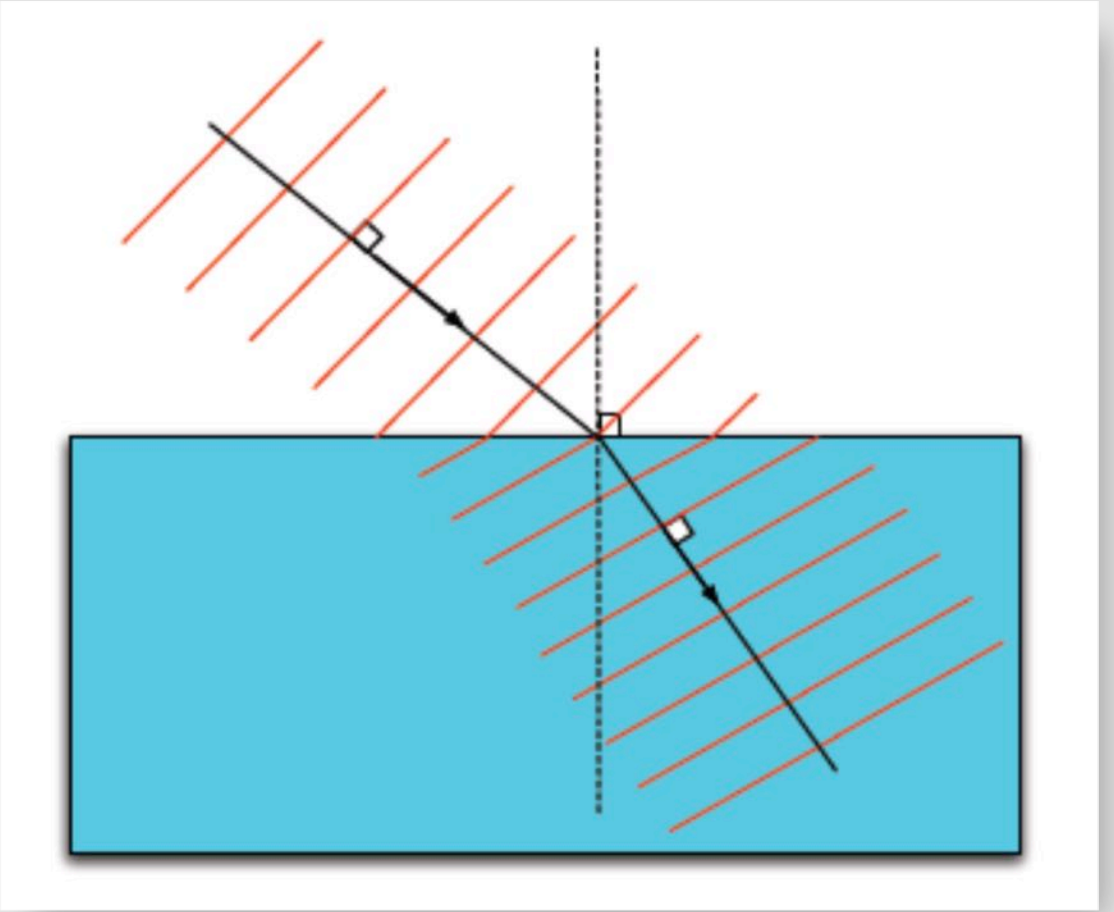
Wavelength dependent bending

George-Alex Koulieris

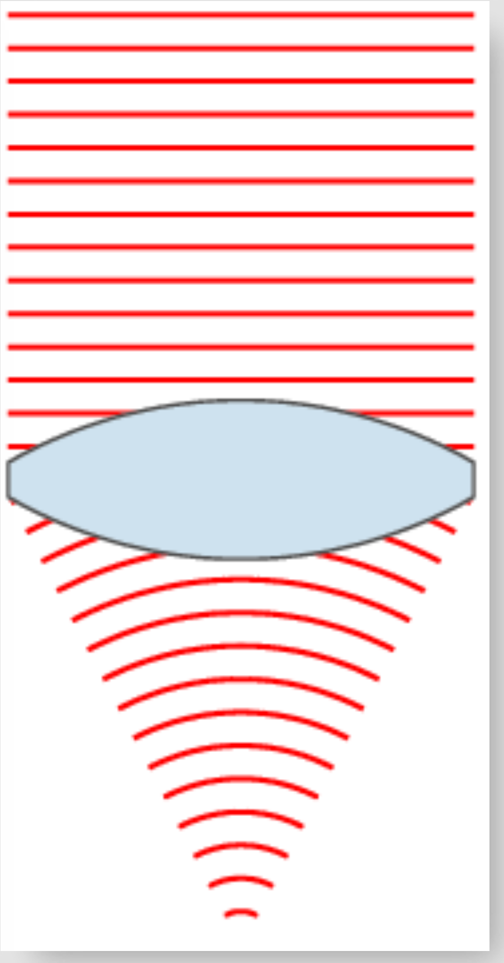


Lucas Barbosa

Light wave-front interacting with a lens



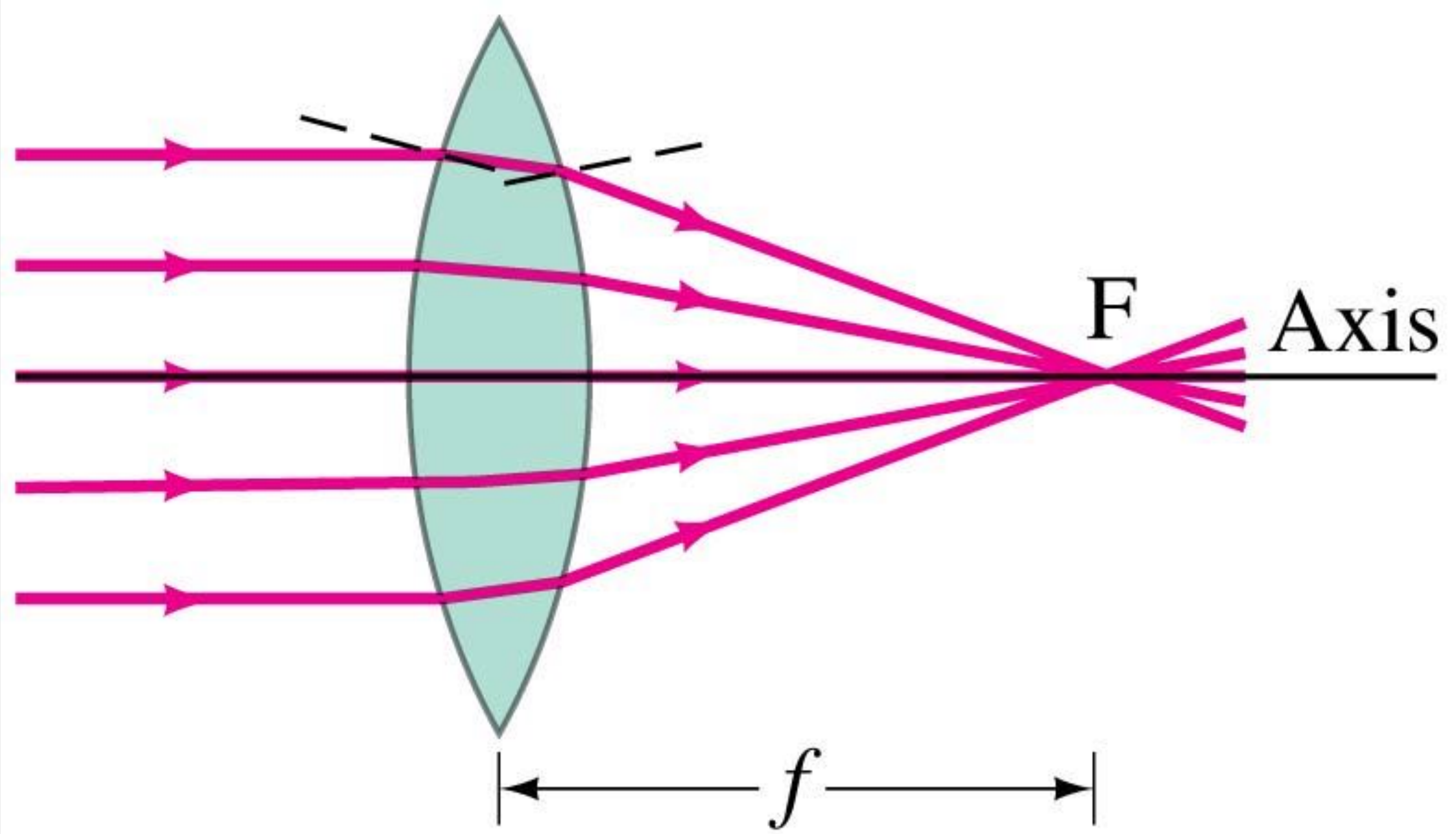
Libretexts



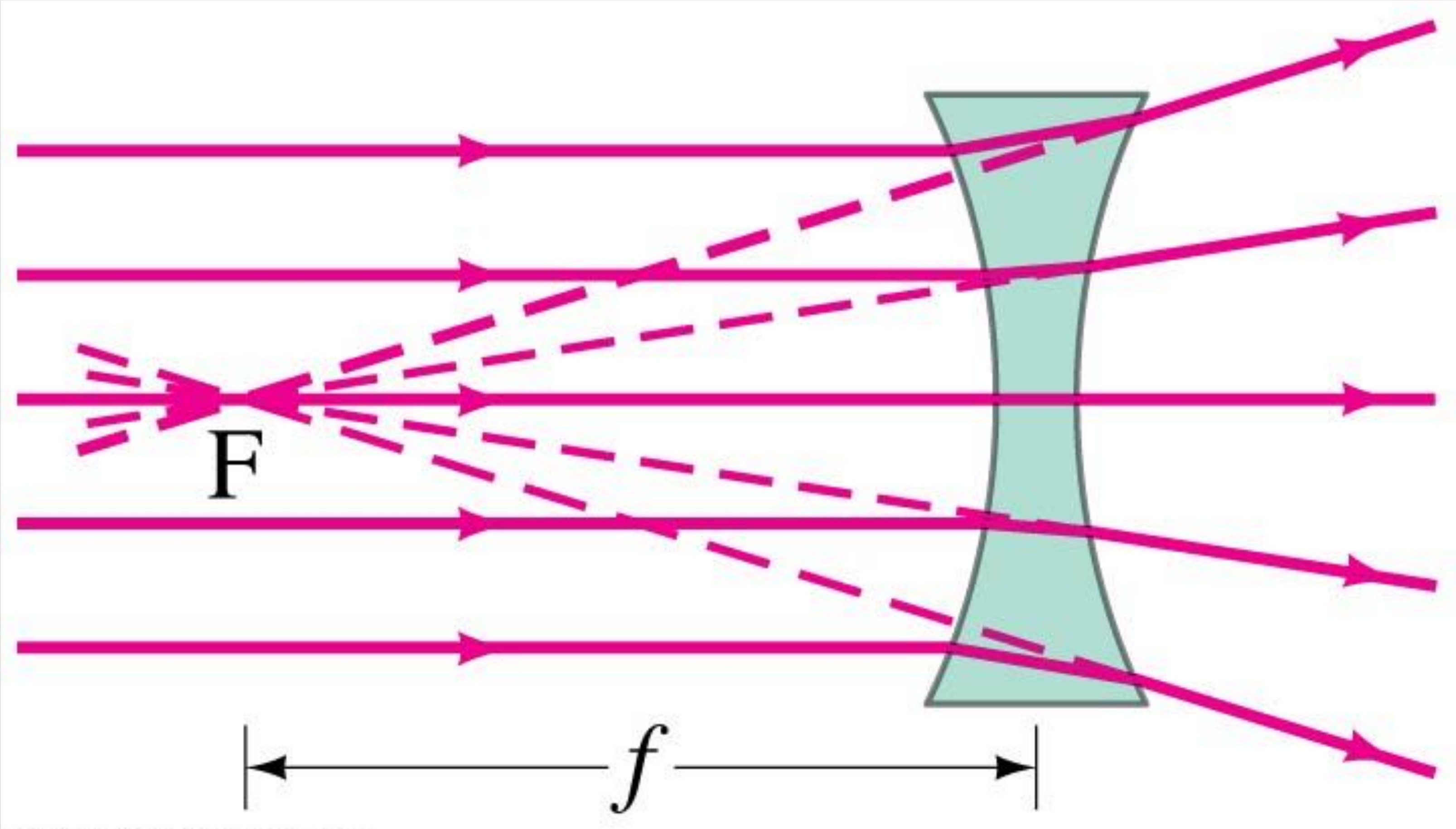
Oleg Alexandrov

George-Alex Koulieris

Convex thin lenses



Concave thin lenses



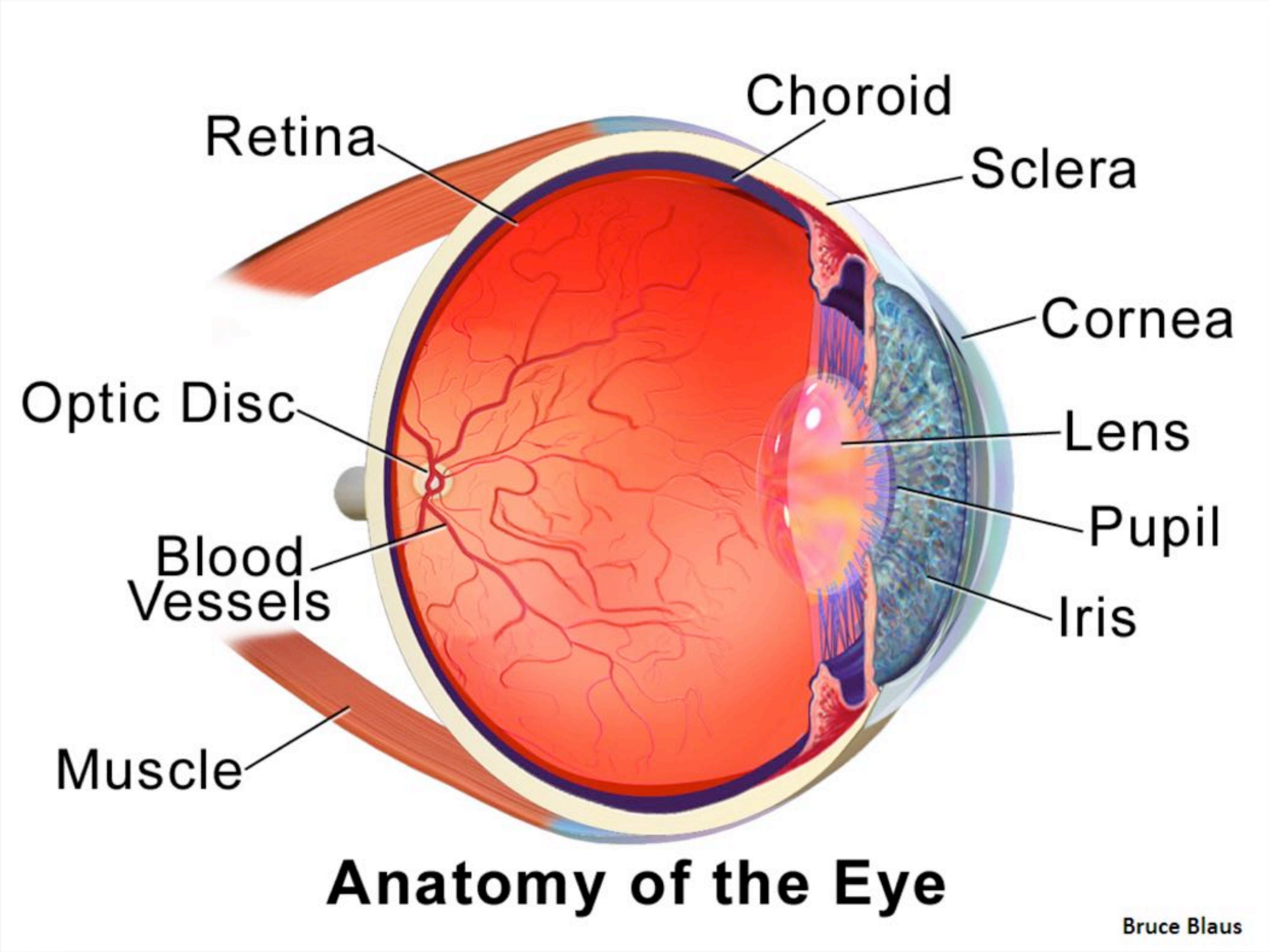
George-Alex Koulieris

Dioptres

- Measurement unit of optical power
- Equal to the reciprocal of the focal length (in m)
- E.g., a 2-dioptre lens brings parallel rays of light to focus at 1 / 2 meter.
- E.g., a flat window has optical power of 0-dioptres
  - does not converge or diverge light.

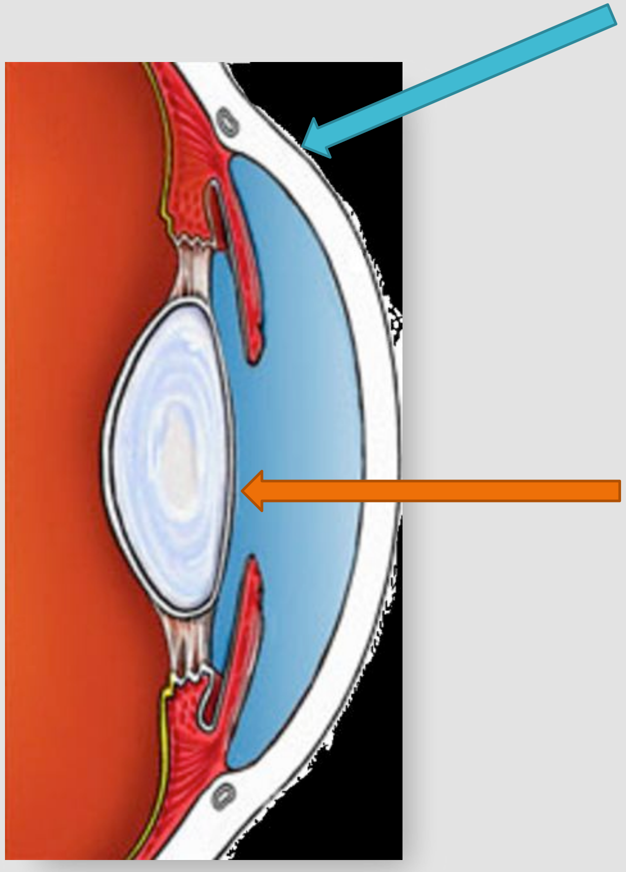
George-Alex Koulieris

Anatomy of the eye



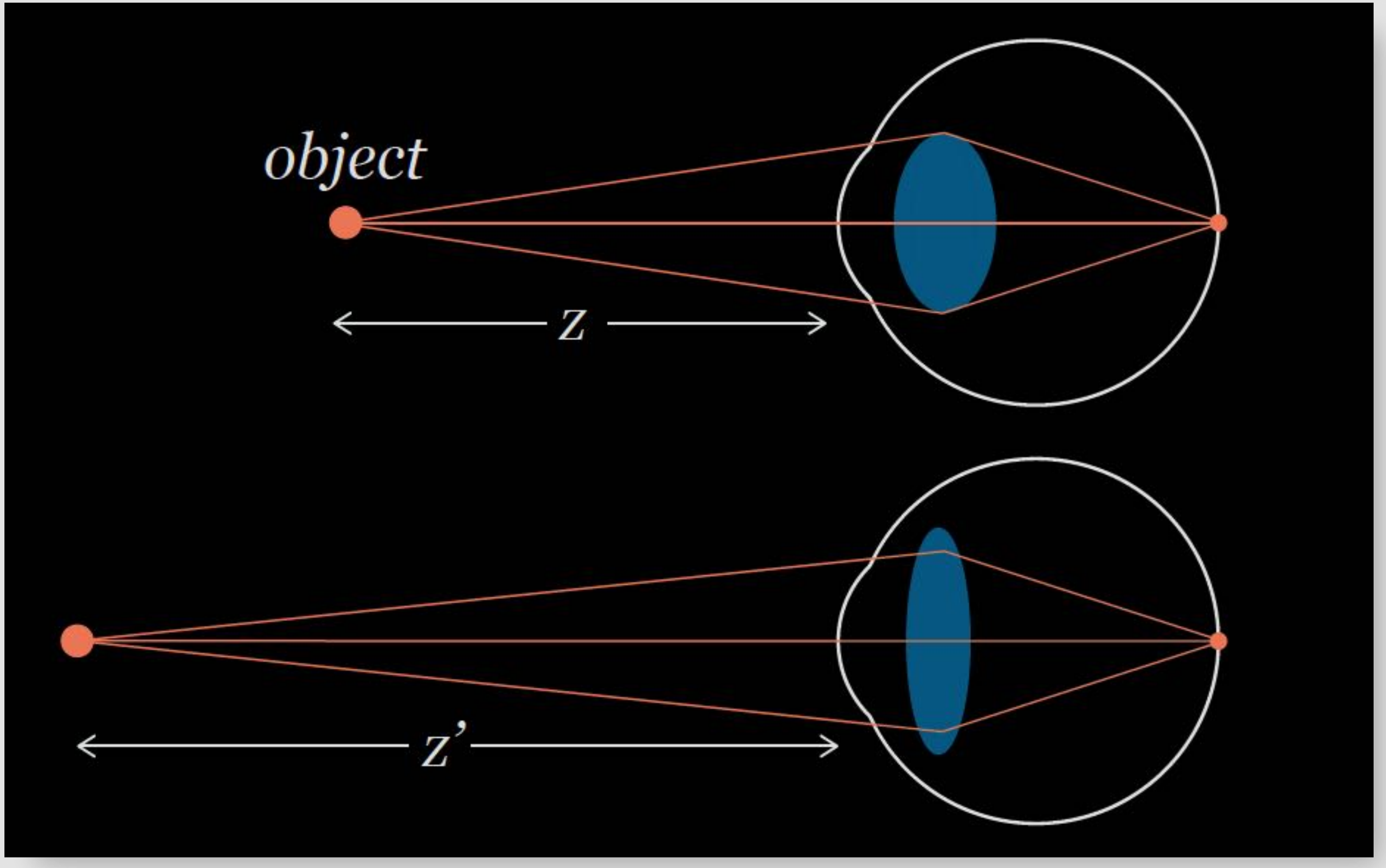
George-Alex Koulieris

Two lenses in the eye



- Cornea
  - **fixed** power ~40 diopters
  - does most of the focusing
  - fun fact: focal length ≈ length of the eye
- Crystalline lens
  - **variable** up to ~20 diopters
  - power diminishes with age (presbyopia)
  - ~350 ms to change power

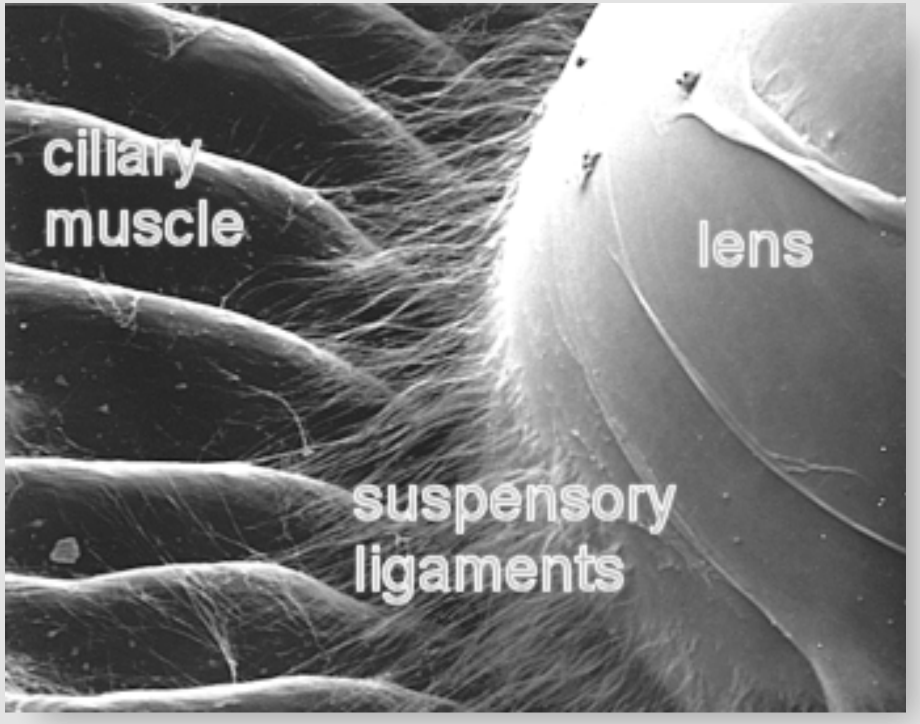
Accommodation



George-Alex Koulieris

Fun fact: accommodation theories

George-Alex Koulieris

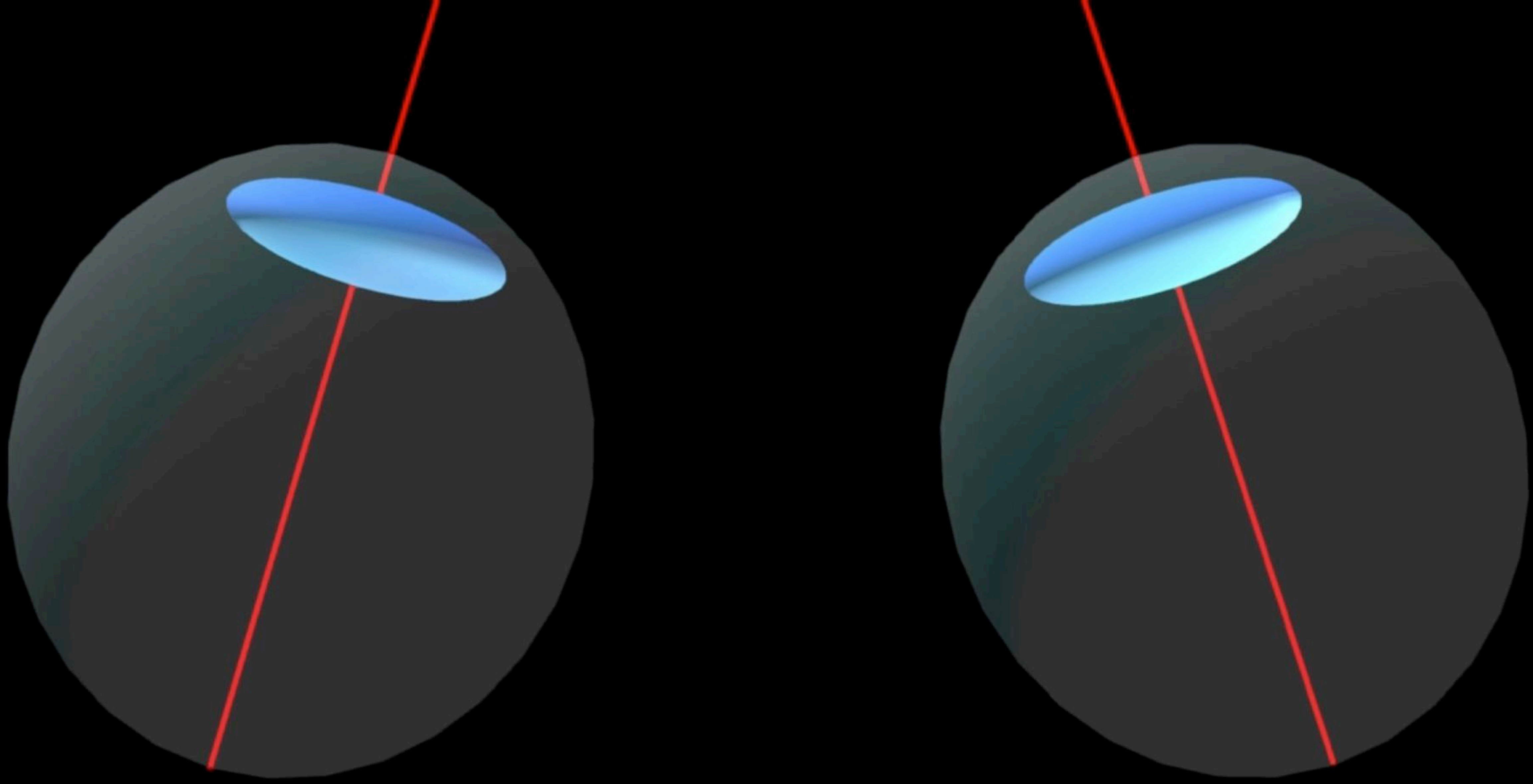


Helmholtz Theory

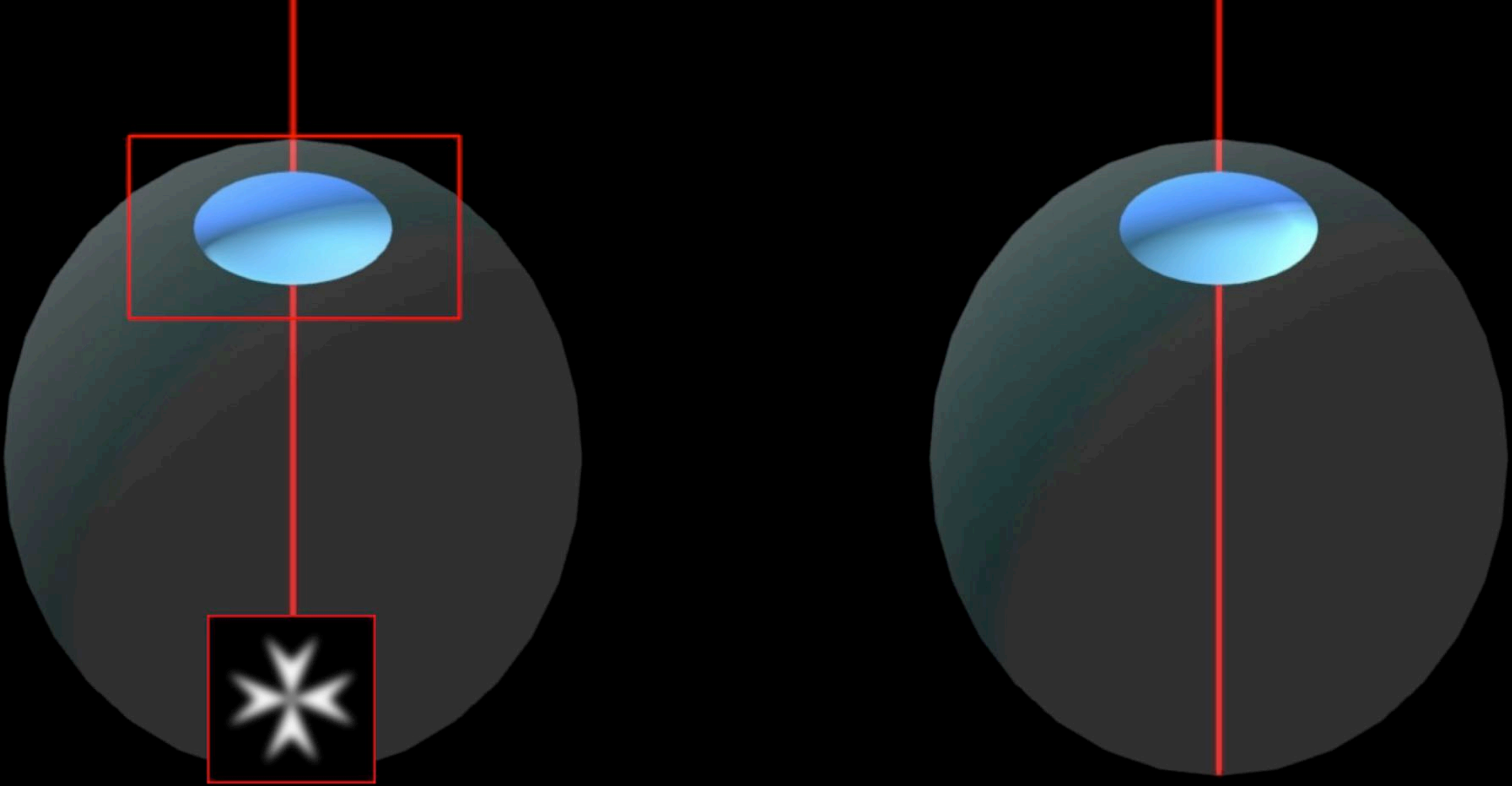
When fixating a near object:

1. circularly arranged ciliary muscle contracts
2. lens zonules and suspensory ligaments relax
3. lens thickens

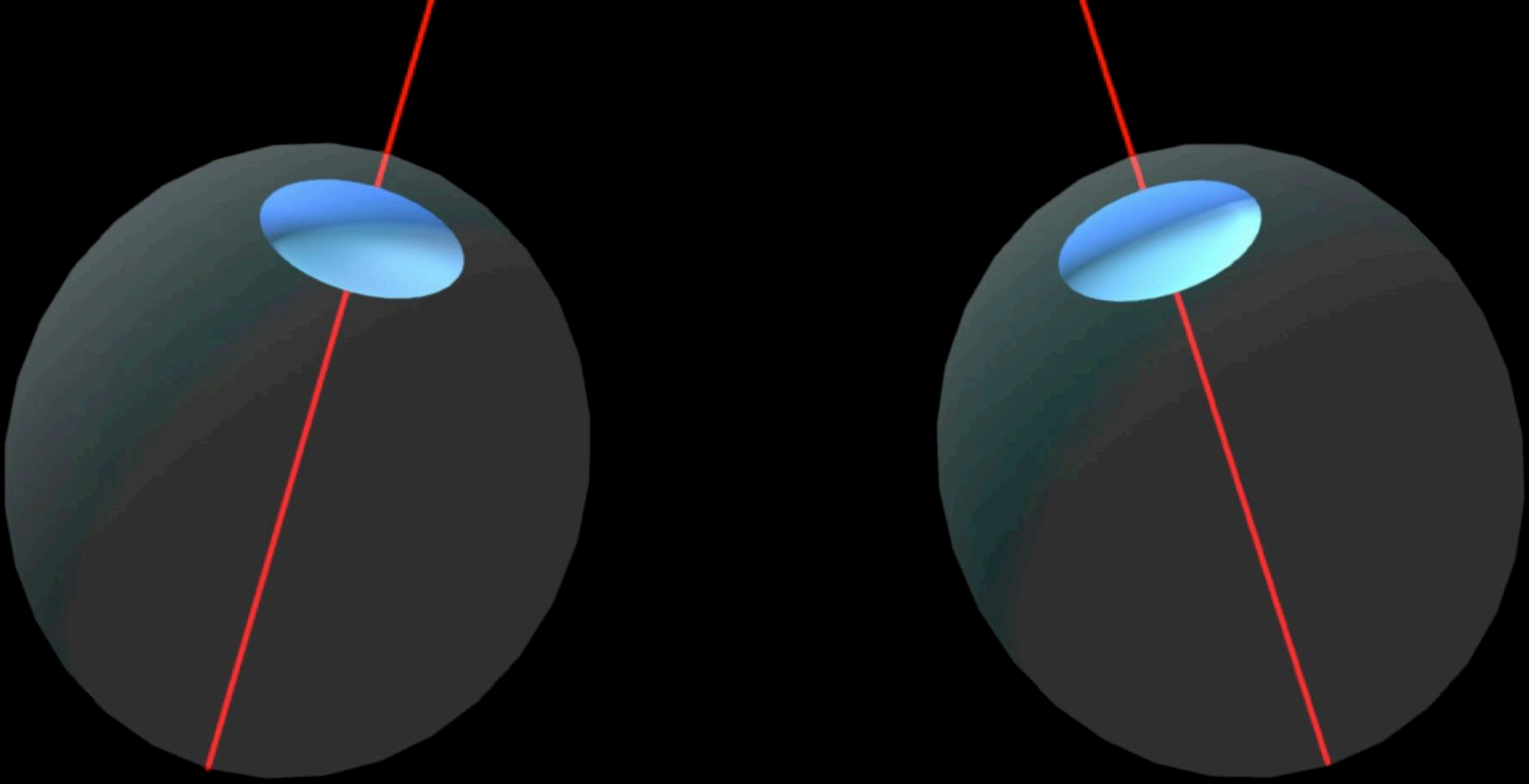
Now let us see how all this relate to a major source of discomfort in VR/AR



VERGENCE

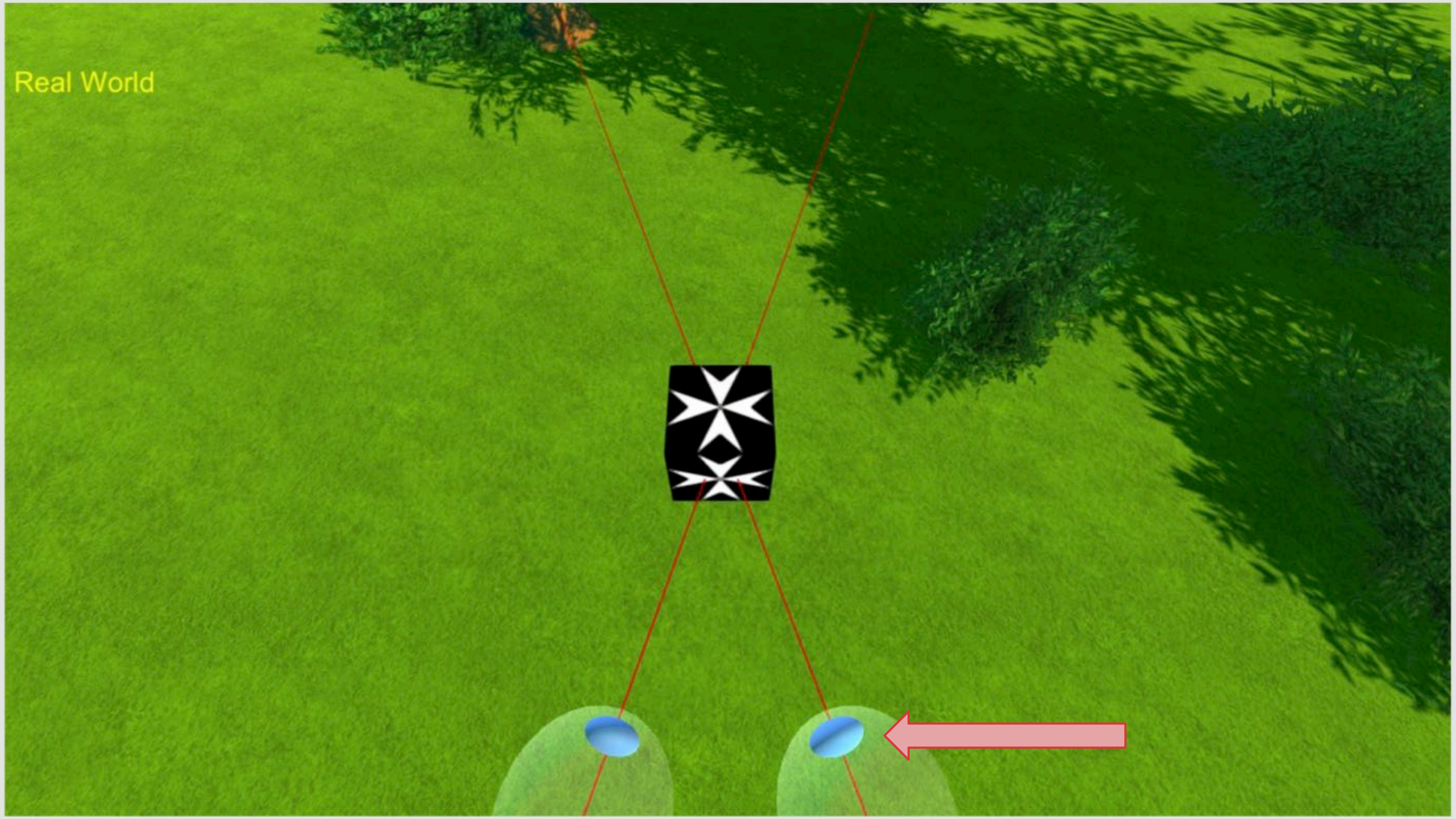


ACCOMMODATION



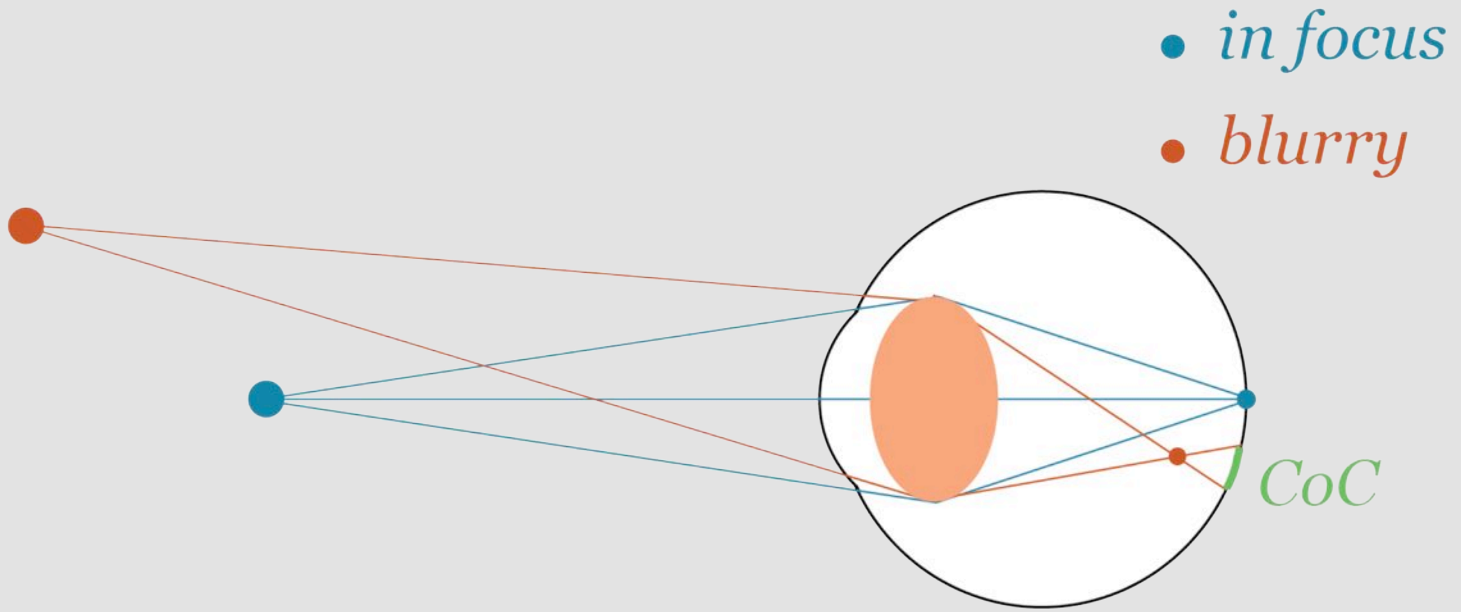
COUPLED

Vergence and accommodation in the real world



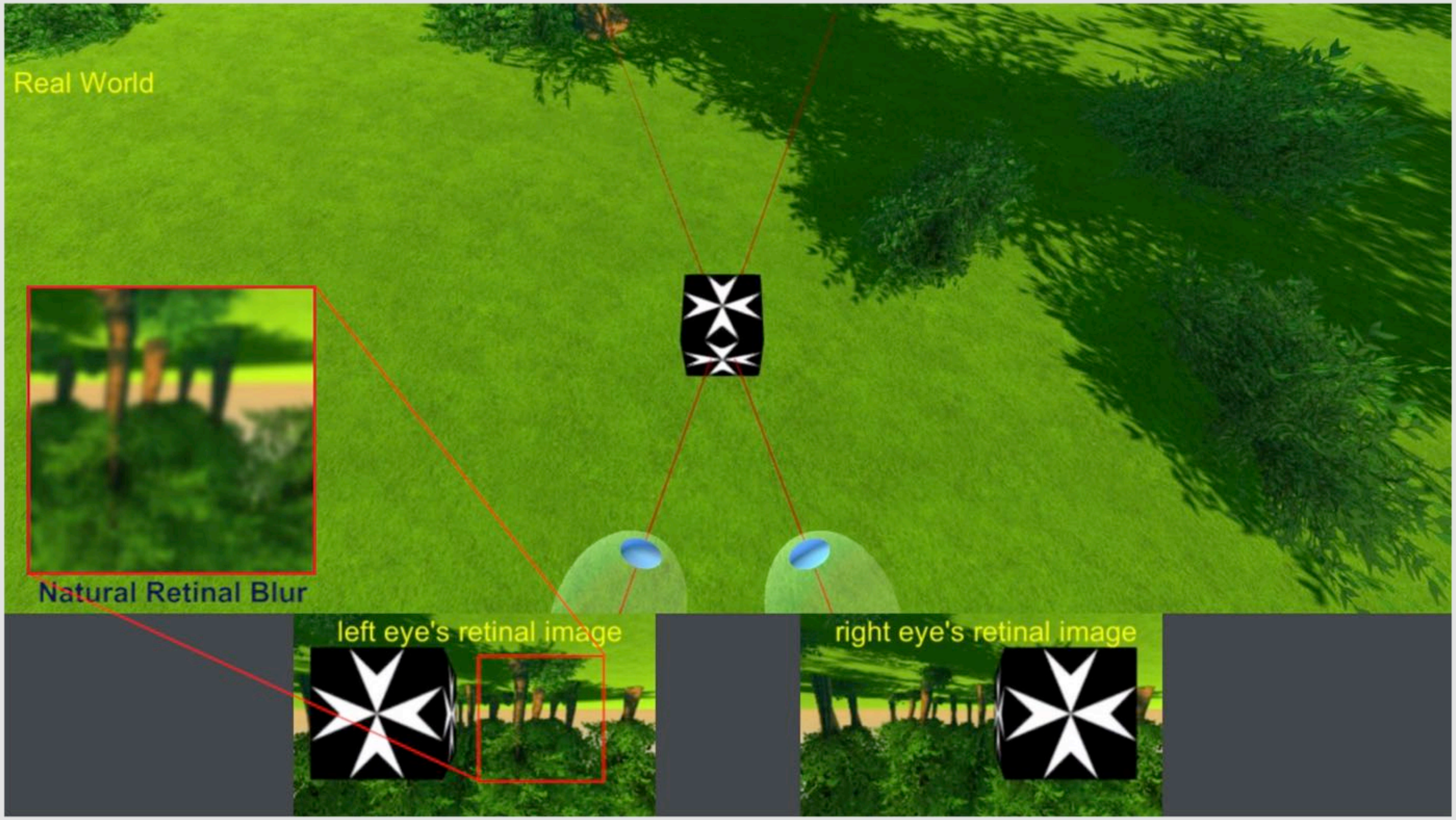
George-Alex Koulieris

Blur in the real world



George-Alex Koulieris

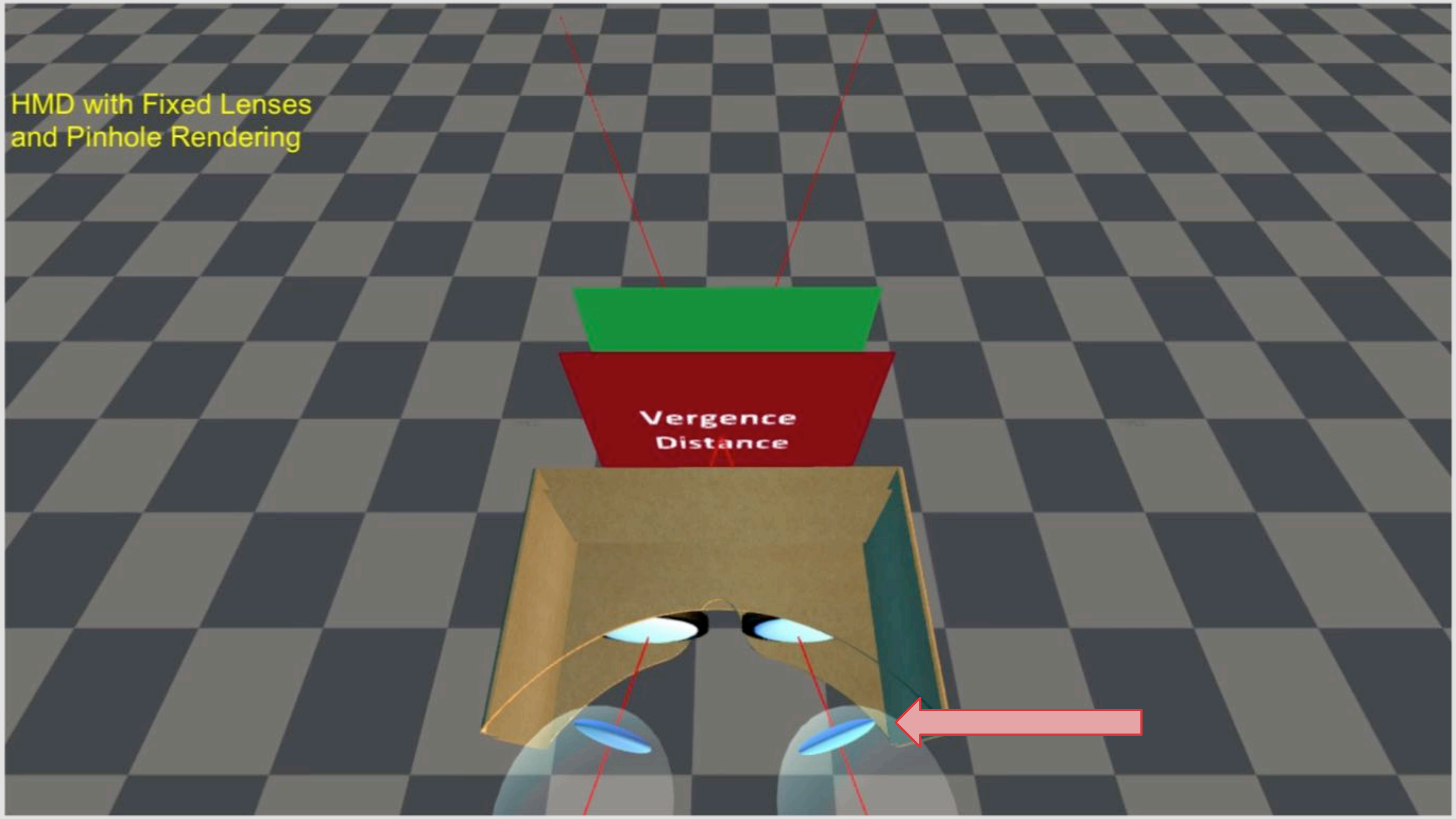
Blur in the real world



George-Alex Koulieris

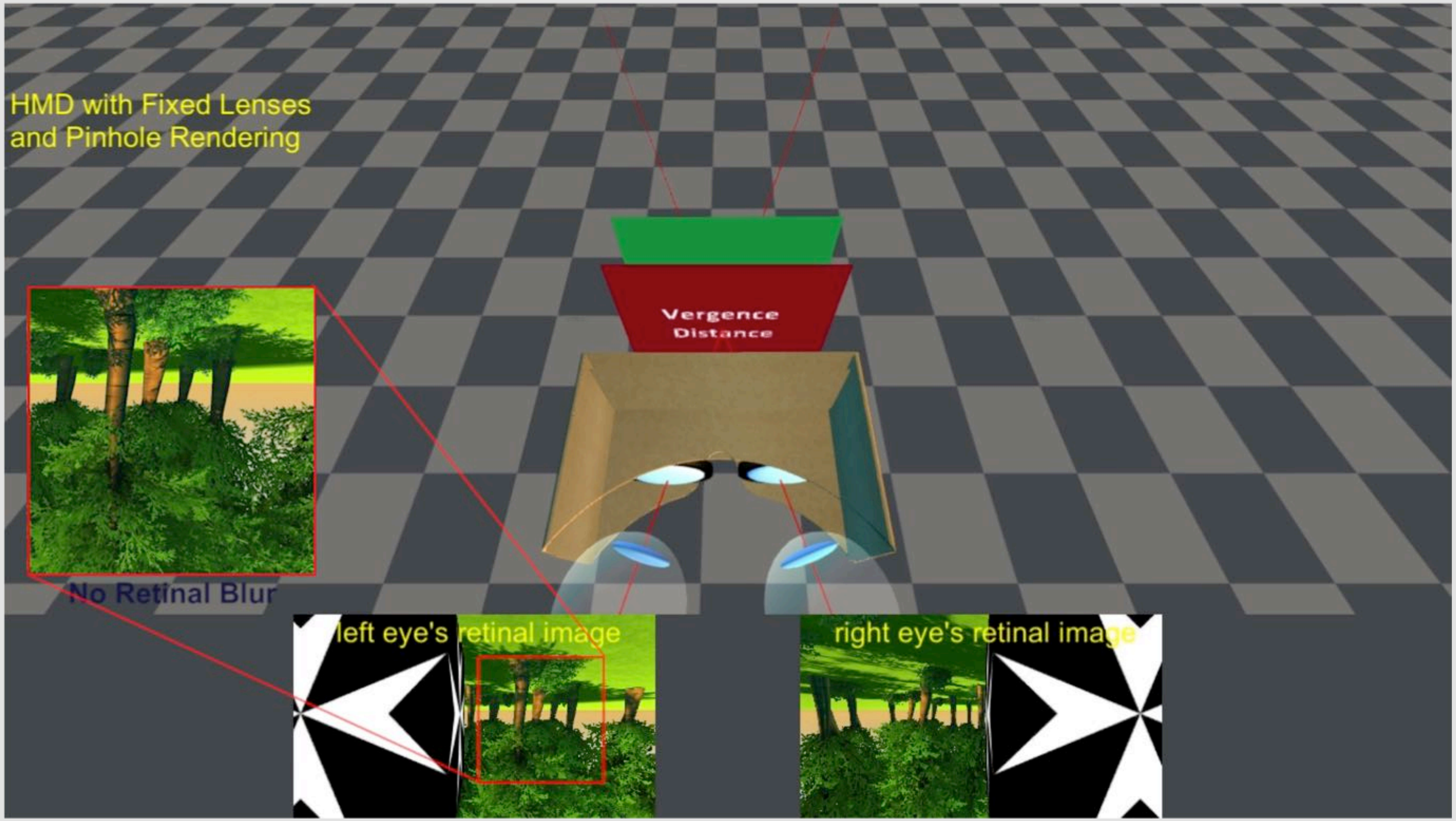
But what about stereo displays ?

Vergence and accommodation conflicting in near-eye displays



George-Alex Koulieris

Blur non-existent in near-eye displays



George-Alex Koulieris

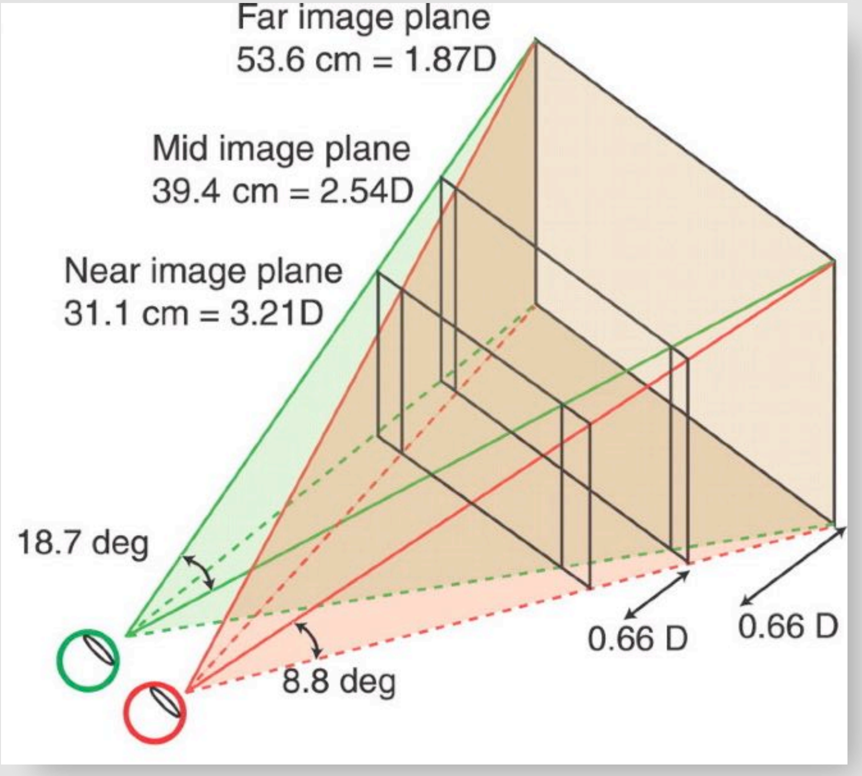
The VA conflict

1. No retinal blur
2. Accommodation generally does not match vergence

Viewer is required to fight against the natural coupling between accommodation and vergence which causes discomfort

George-Alex Koulieris

VA conflict is a major source of discomfort

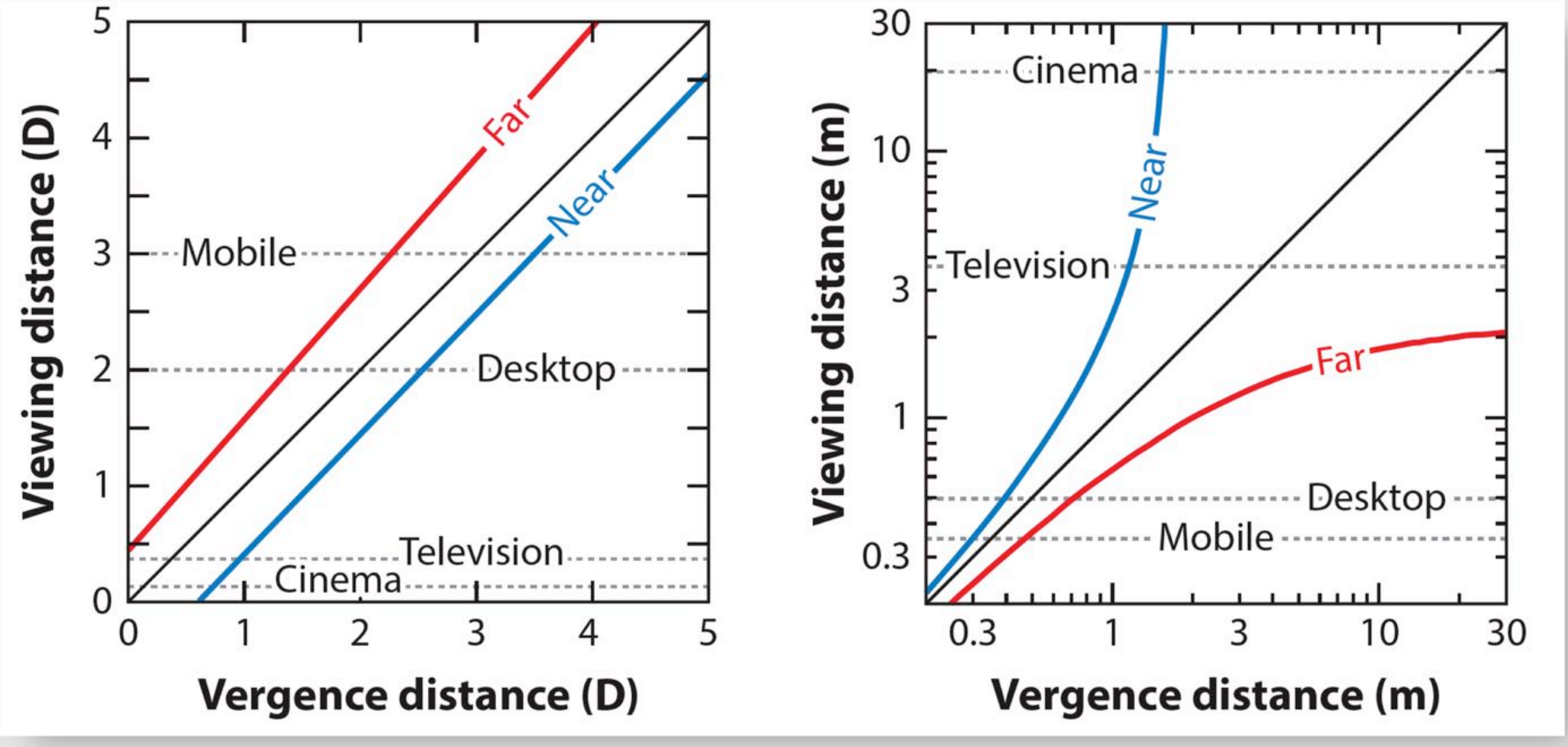


Hoffman & Banks, 2010

Hoffman, D. M., Girshick, A. R., Akeley, K., & Banks, M. S. (2008). Vergence–accommodation conflicts hinder visual performance and cause visual fatigue. *Journal of vision*, *8*(3), 33-33.

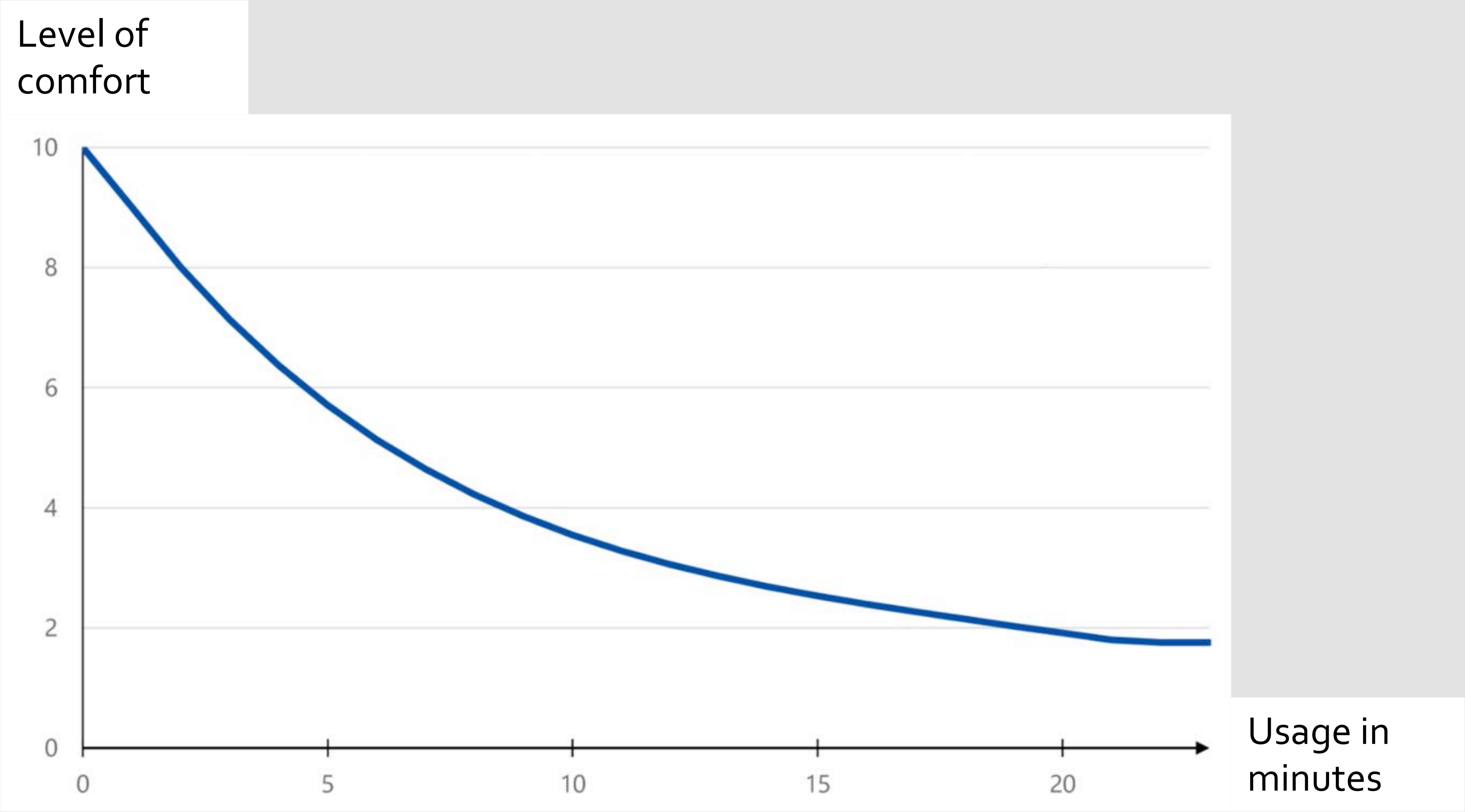
George-Alex Koulieris

VA conflict is a major source of discomfort



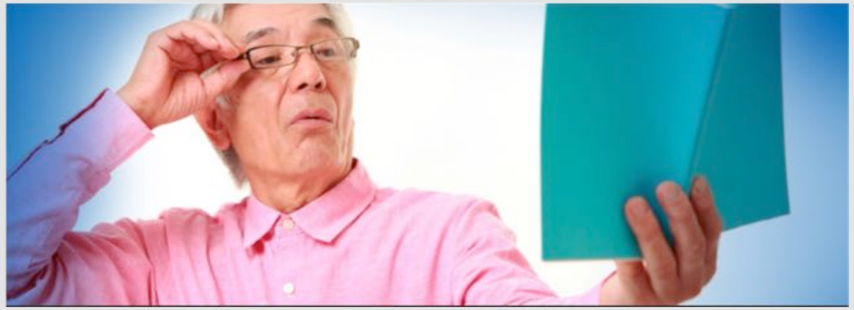
Shibata, T., Kim, J., Hoffman, D. M., & Banks, M. S. (2011). The zone of comfort: Predicting visual discomfort with stereo displays. *Journal of vision*, *11*(8), 11-11.

Comfort in VR, today



Fernandes and Feiner, 2016

VA conflict in presbyopes



eye-trends.com

- Range of distances one can accommodate declines starting at the age of 40
- By 50/60 accommodative range is essentially zero
- Presbyopes are always in conflict ➔ used to it!
- No VA conflict due to stereoscopic viewing

Yang, S. N., Schlieski, T., Selmins, B., Cooper, S. C., Doherty, R. A., Corriveau, P. J., & Sheedy, J. E. (2012). Stereoscopic viewing and reported perceived immersion and symptoms. *Optometry and vision science*, *89*(7), 1068-1080.

George-Alex Koulieris

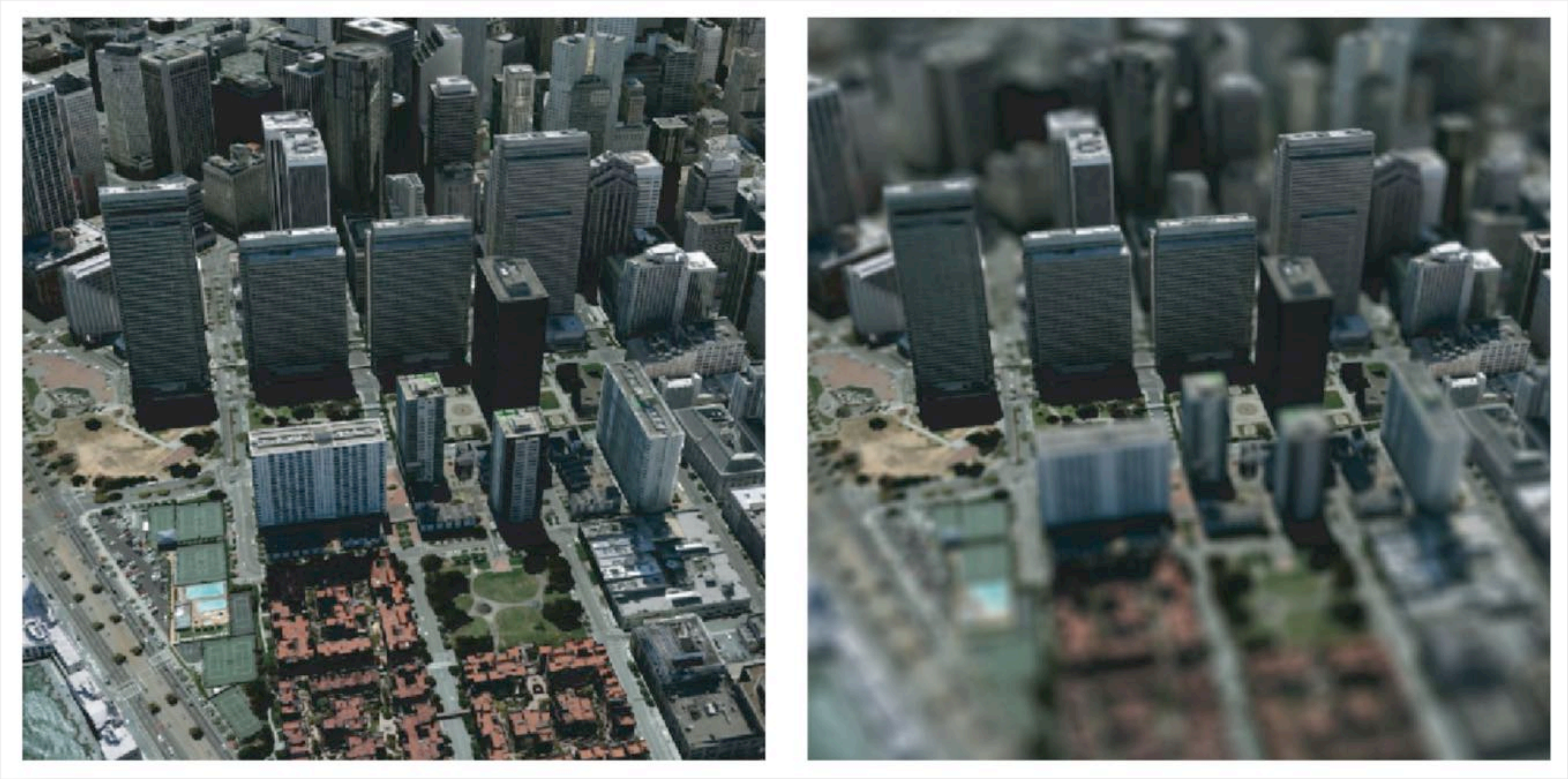
Lack of focus cues is not only affecting discomfort

- 3D shape perception
- Apparent scale of scenes
- Binocular performance

Buckley, D., & Frisby, J. P. (1993). Interaction of stereo, texture and outline cues in the shape perception of three-dimensional ridges. *Vision research*, *33*(7), 919-933.

Watt, S. J., Akeley, K., Ernst, M. O., & Banks, M. S. (2005). Focus cues affect perceived depth. *Journal of vision*, *5*(10), 7-7.

Focus cues affect perceived size of scenes



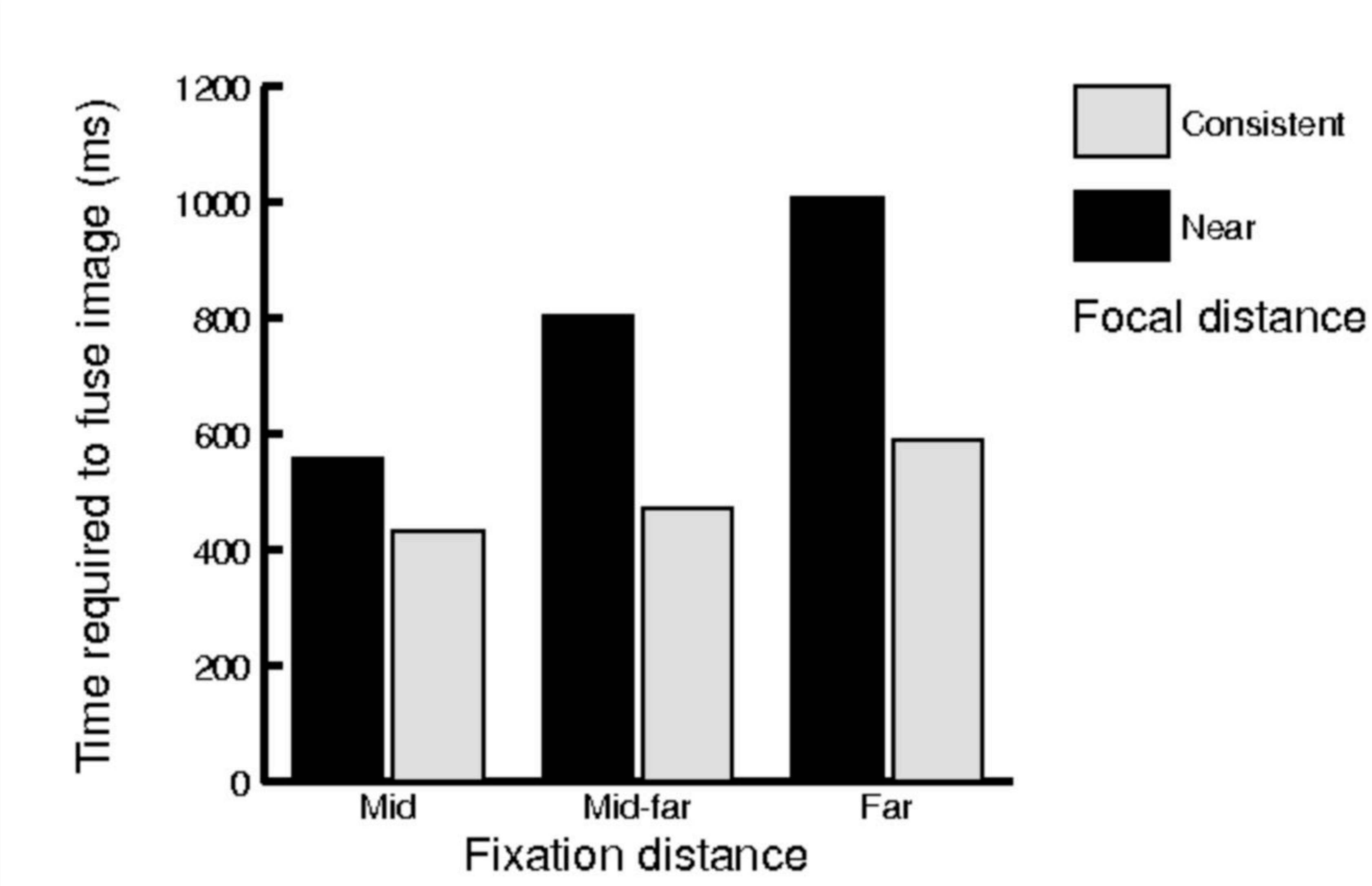
Held et al., 2010

Fielding R. 1985. Techniques of Special Effects Cinematography. Oxford, UK: Focal Press. 4th ed.

Held RT, Cooper EA, O'Brien JF, Banks MS. 2010. Using blur to affect perceived distance and size. ACM Trans. Graph. 29(2):19

George-Alex Koulieris

Focus cues affect visual performance



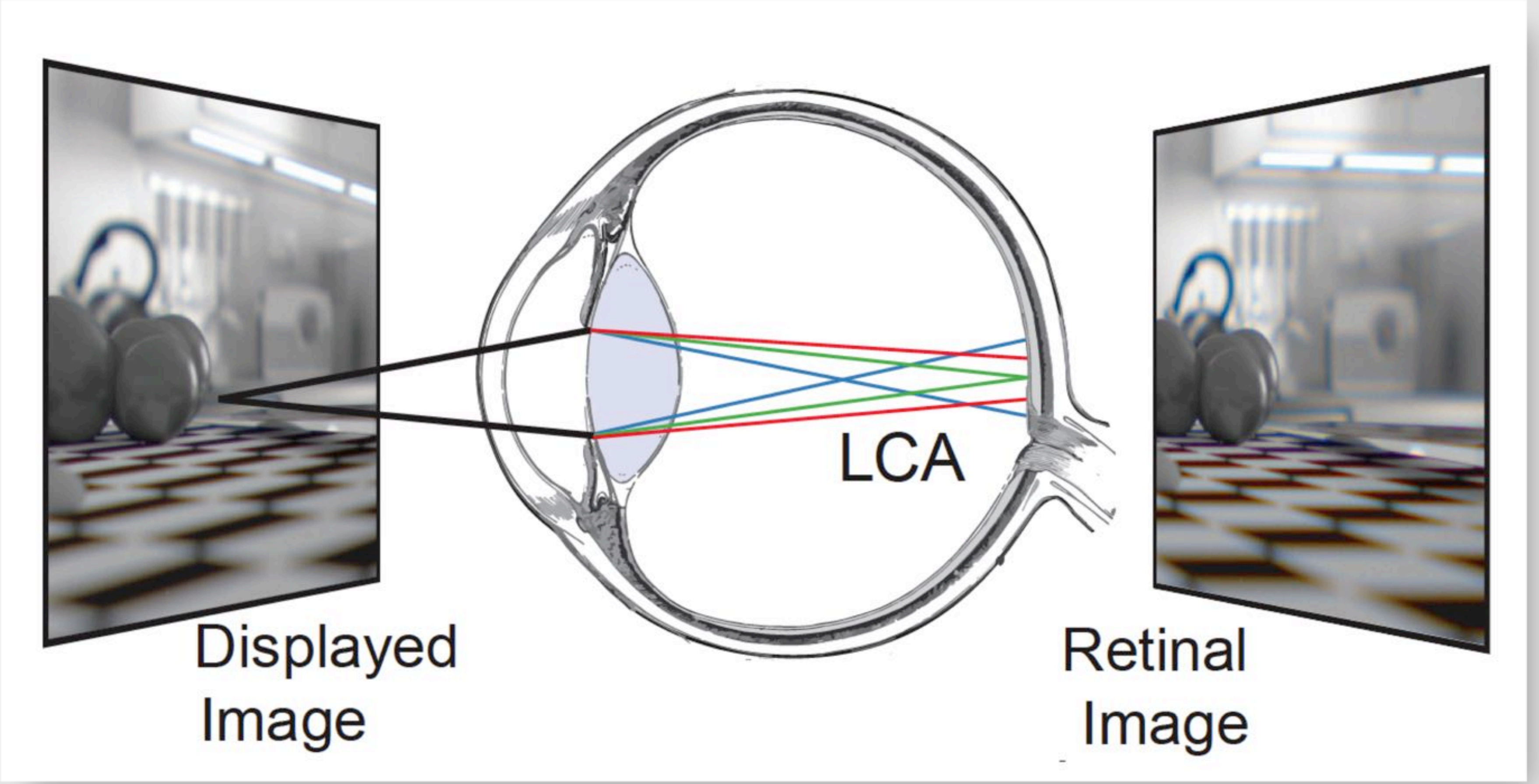
Akeley, K., Watt, S. J., Girshick, A. R., & Banks, M. S. (2004, August). A stereo display prototype with multiple focal distances. In *ACM transactions on graphics (TOG)* (Vol. 23, No. 3, pp. 804-813). ACM.

George-Alex Koulieris

Fun facts: the iris

1. Reduces light by a factor of ~20
2. Constriction increases depth-of-field
3. Reduces spherical aberration by occluding outer parts of lens

Fun fact: accommodation and the retina

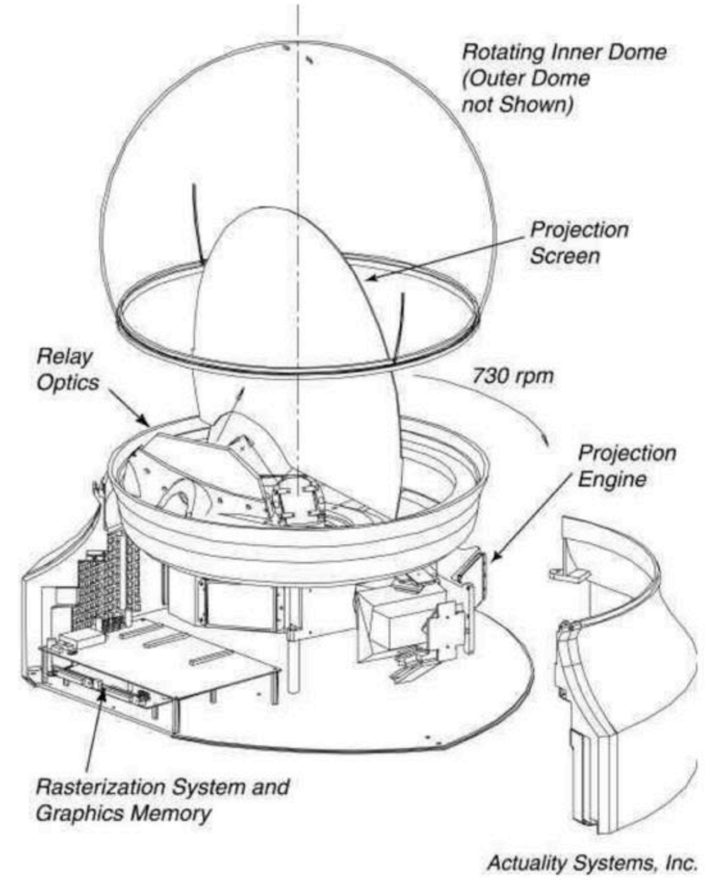


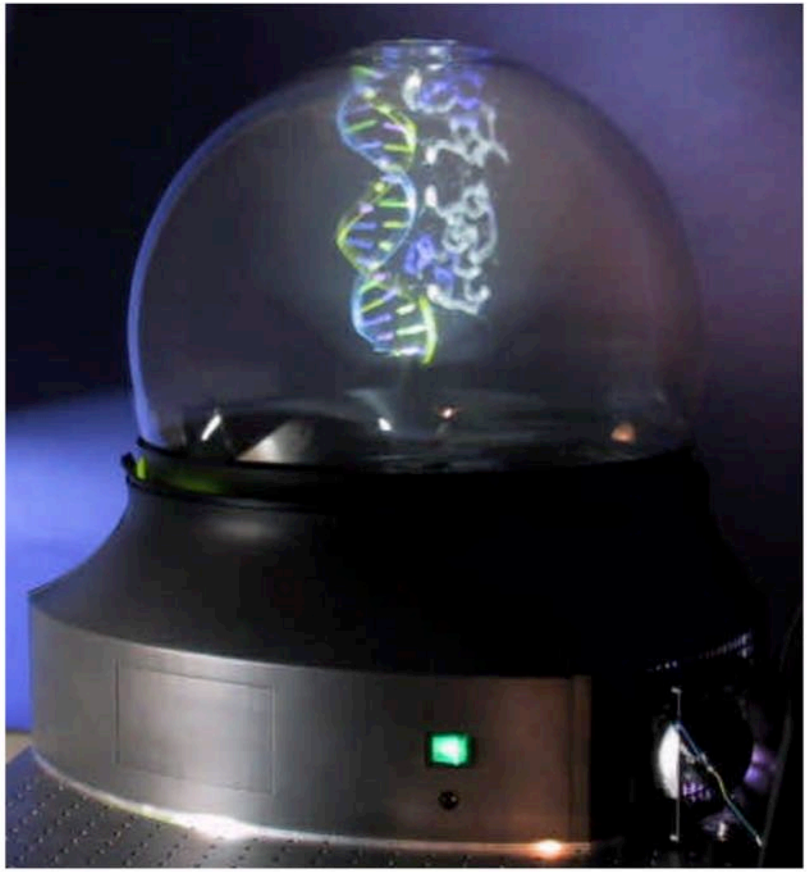
adapted from Cholewiak et al., 2017

- Infinitesimal amount of S-Cones ("blue") in the fovea
  - due to Longitudinal Chromatic Aberration?

Part 2:
Multifocal displays

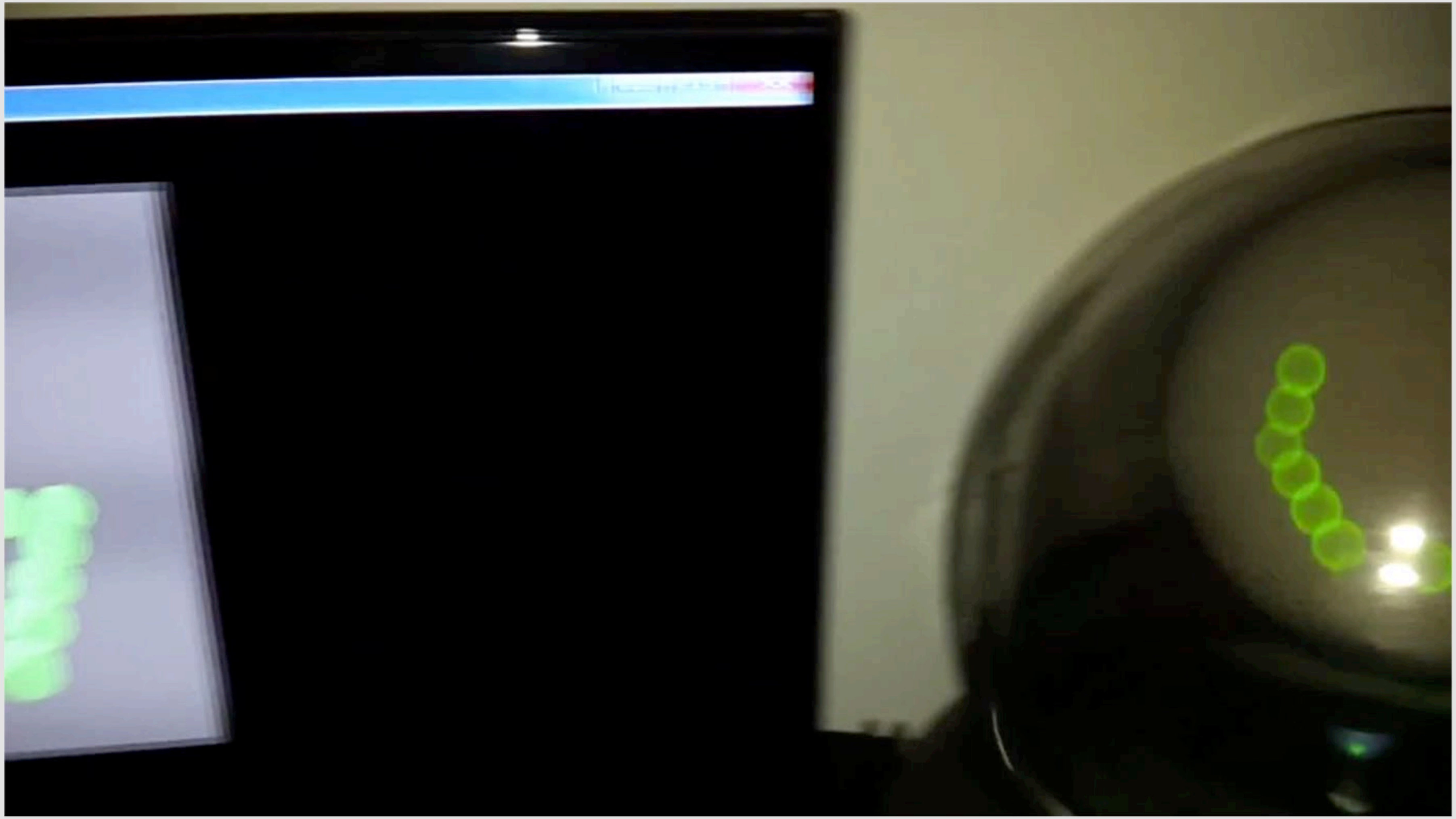
Swept-screen volumetric displays





Favalora, G. E., Napoli, J., Hall, D. M., Dorval, R. K., Giovinco, M., Richmond, M. J., & Chun, W. S. (2002, August). 100-million-voxel volumetric display. In *Cockpit Displays IX: Displays for Defense Applications* (Vol. 4712, pp. 300-313). International Society for Optics and Photonics.

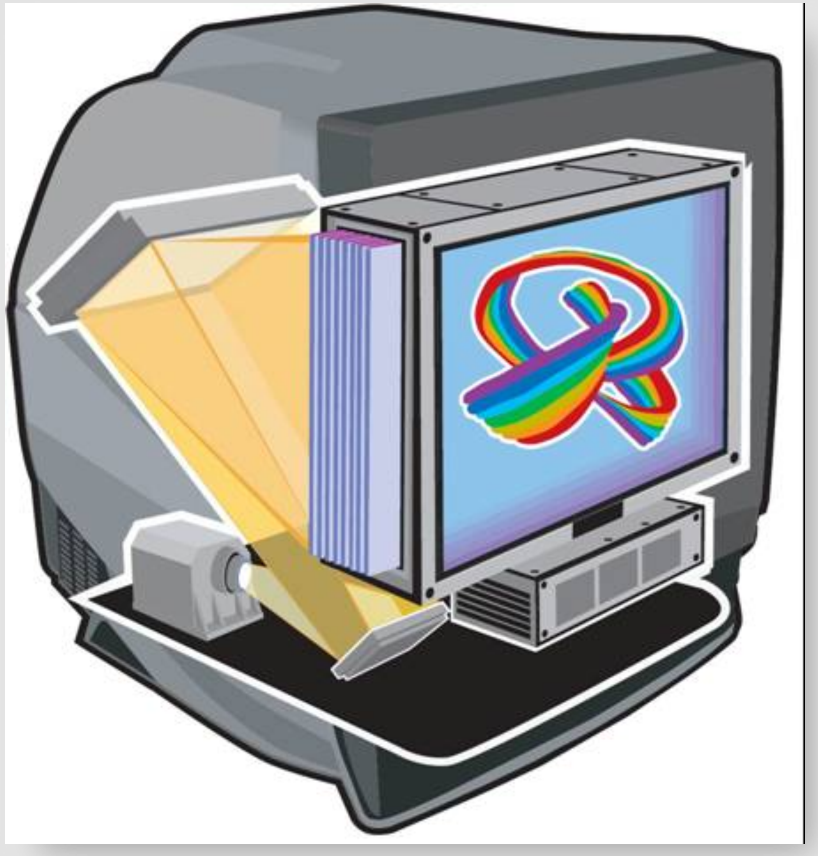
Swept-screen volumetric displays



Abhijit Karnik

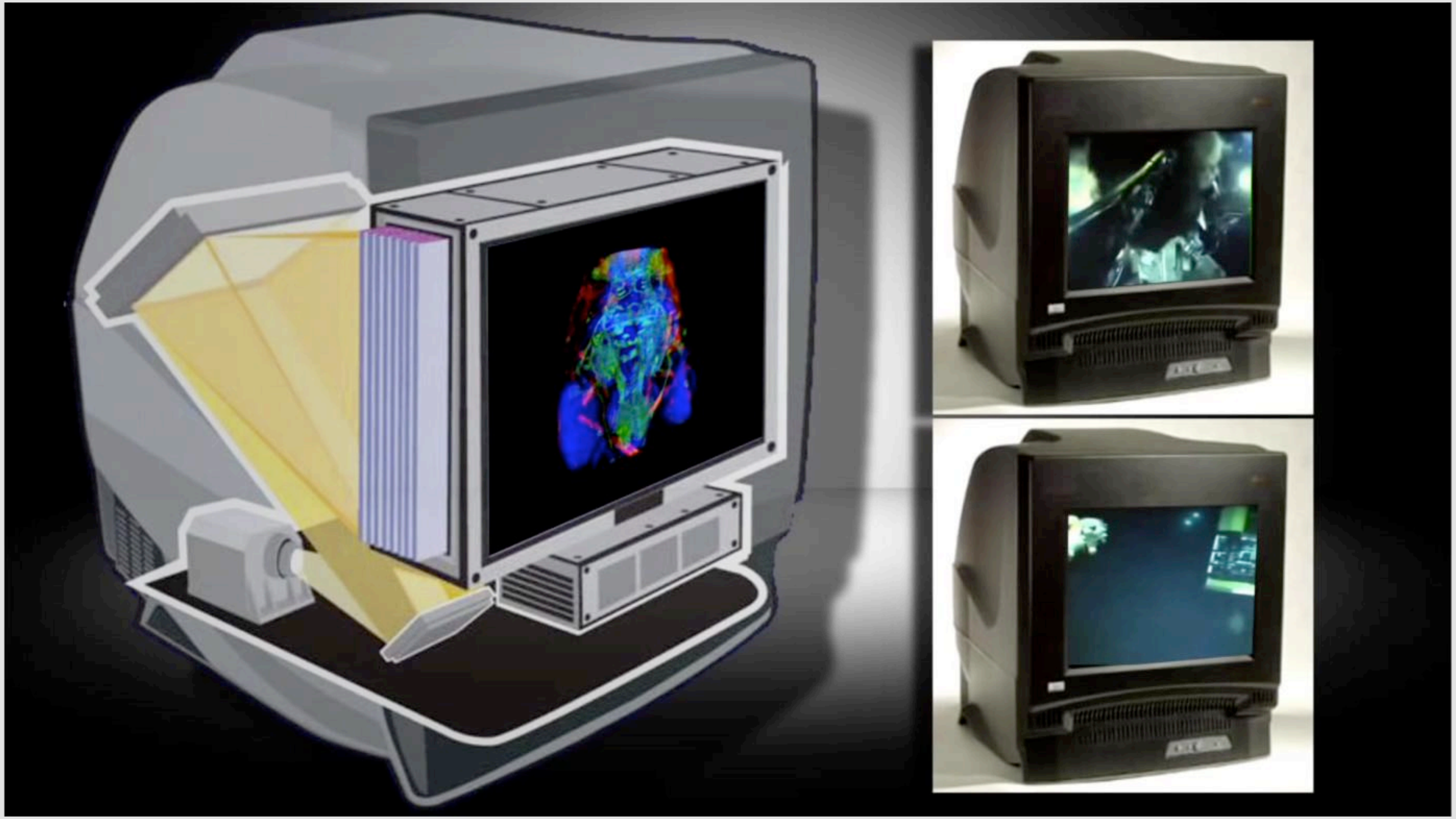
George-Alex Koulieris

Stacked-screen volumetric displays



Sullivan, A. (2004, May). DepthCube solid-state 3D volumetric display. In *Stereoscopic displays and virtual reality systems XI*(Vol. 5291, pp. 279-285). International Society for Optics and Photonics.

Stacked-screen volumetric displays

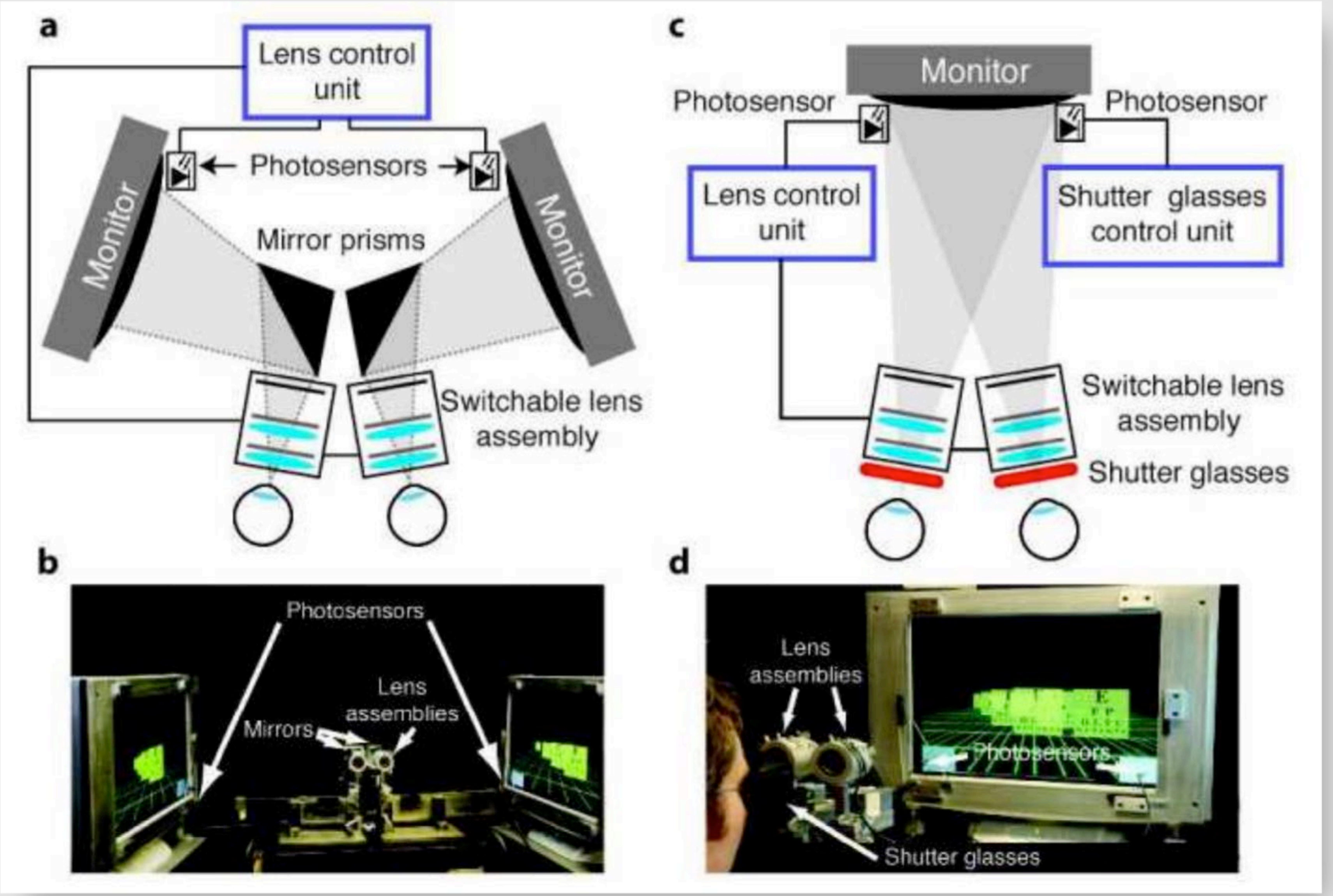


LightSpace Technologies

Advantages & disadvantages

- Present correct stereo, parallax and focus cues
- BUT
- Displayed scene confined to display volume
- Require computing and addressing a huge number of addressable voxels
- Cannot reproduce occlusions and viewpoint-dependent effects (e.g., reflections)

Fixed view-point volumetric displays



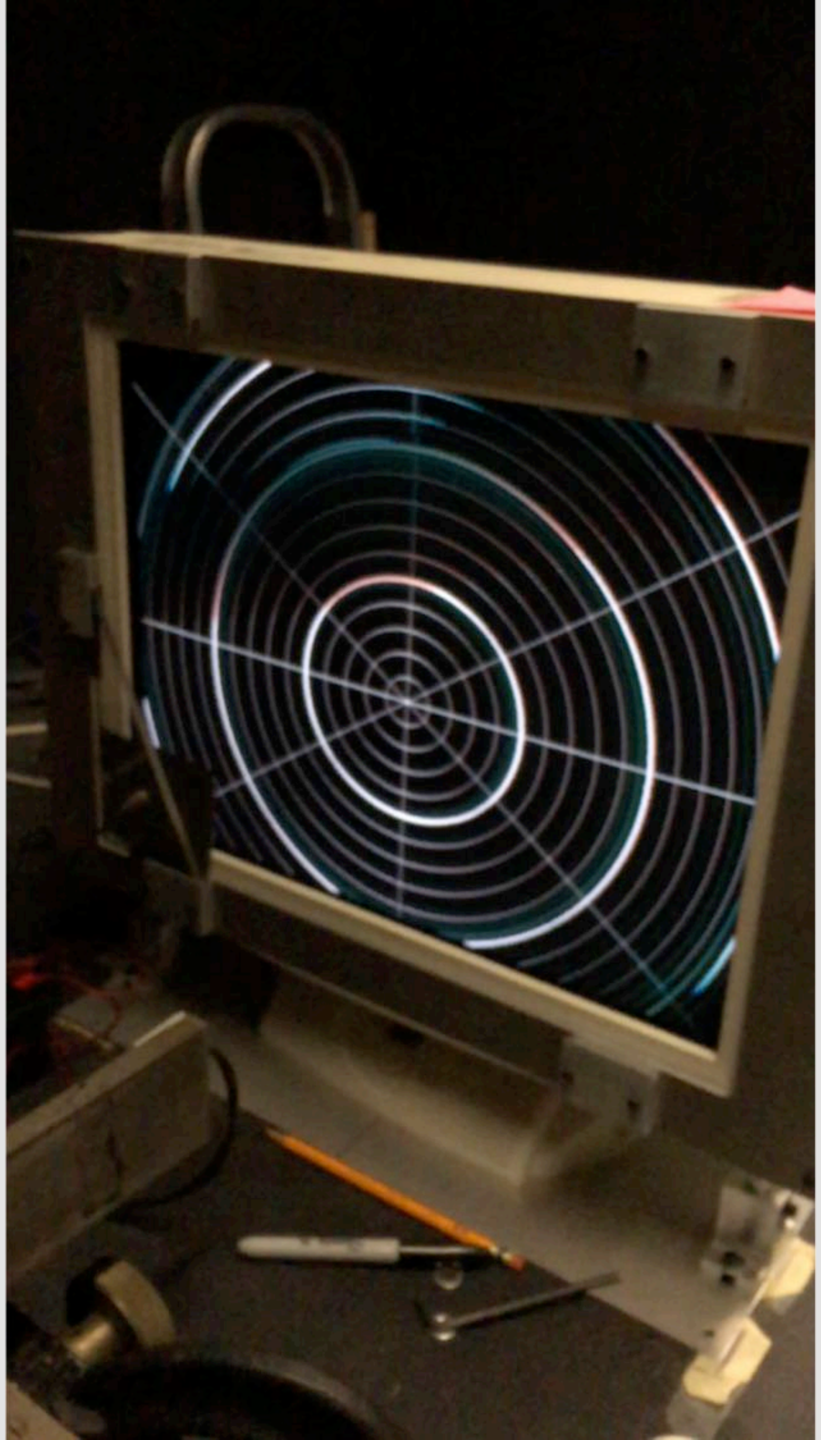
Love, G. D., Hoffman, D. M., Hands, P. J., Gao, J., Kirby, A. K., & Banks, M. S. (2009). High-speed switchable lens enables the development of a volumetric stereoscopic display. *Optics express*, *17*(18), 15716-15725.

Fixed view-point volumetric displays

- Images drawn on presentation planes at different focal distances
- Superimposition of multiple presentation planes additively on the retina
- Special treatment of scene points in between depth planes

Narain, R., Albert, R. A., Bulbul, A., Ward, G. J., Banks, M. S., & O'Brien, J. F. (2015). Optimal presentation of imagery with focus cues on multi-plane displays. *ACM Transactions on Graphics (TOG)*, *34*(4), 59.

Operation



George-Alex Koulieris

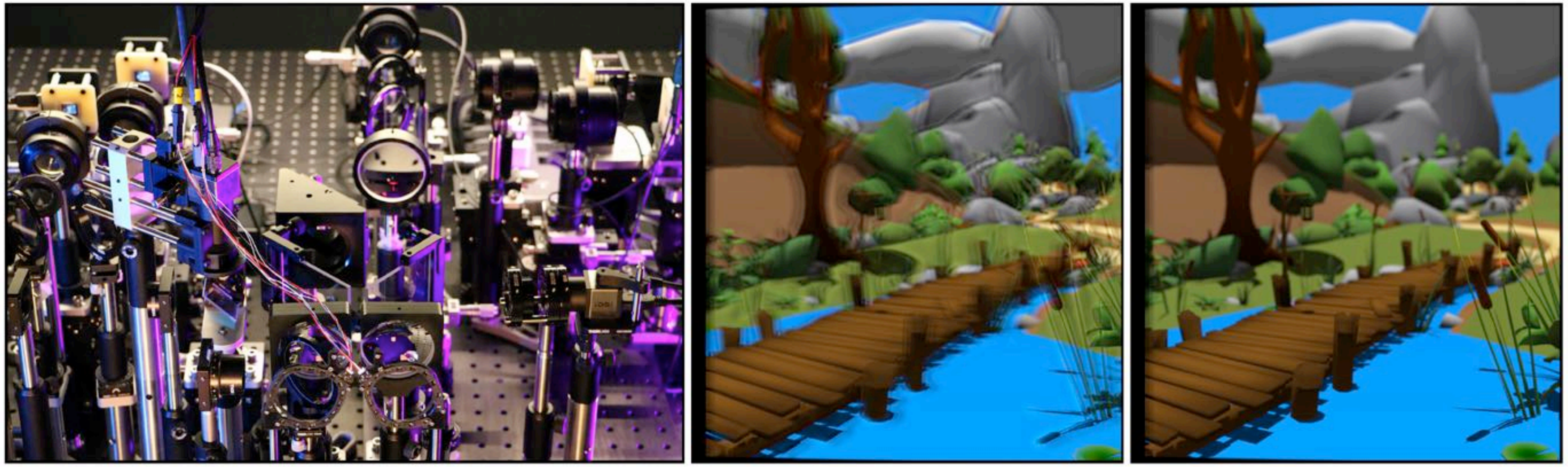
Advantages & disadvantages

- Very high resolution
- Accommodation cues
- Comfortable

- BUT

- Need to fixate head using bite-bars or other means

Fast gaze-contingent decomposition for multifocal displays

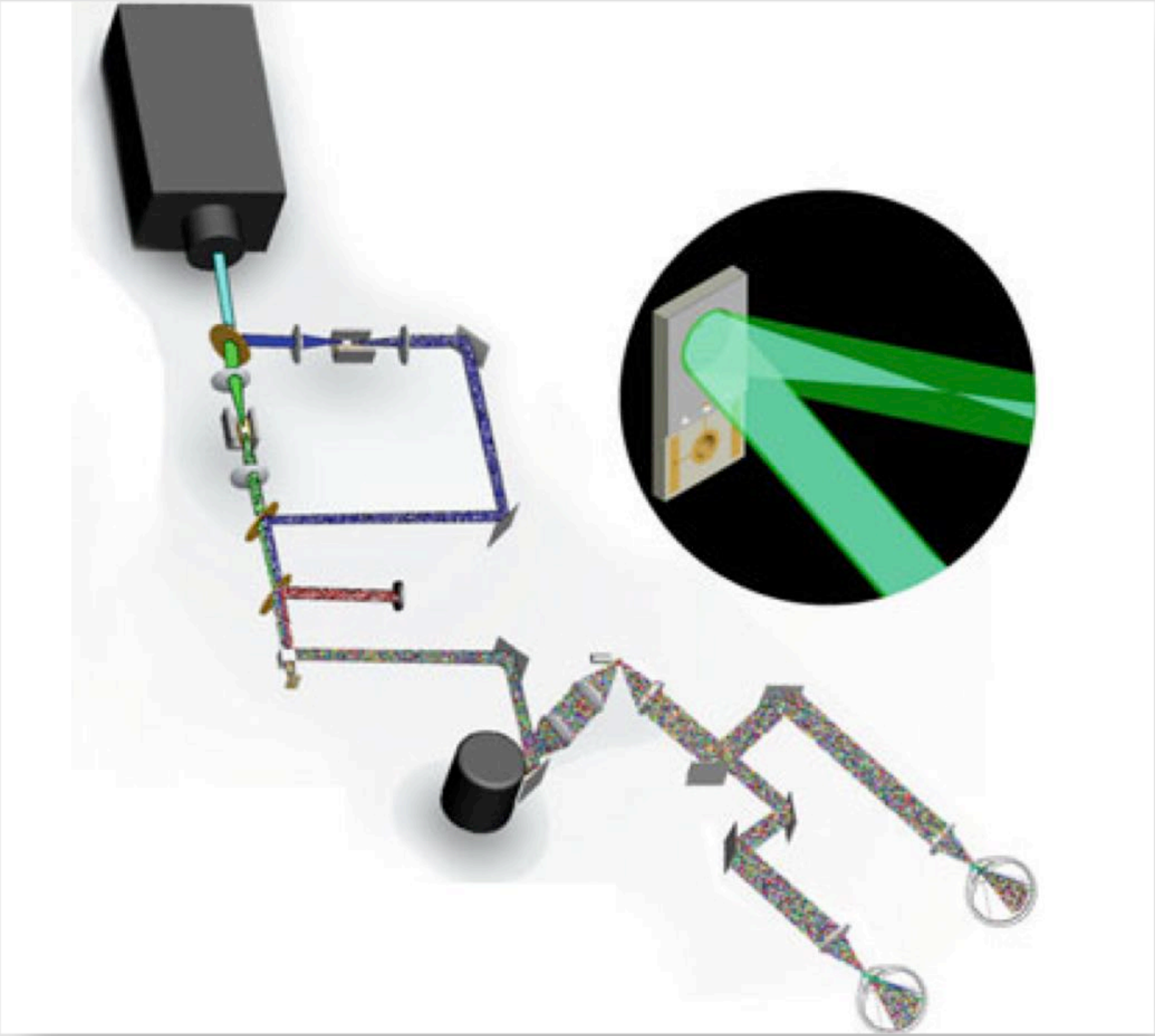


(a) Multifocal Testbed with Eye and Accommodation Tracking (b) Eye Movement without Correction (c) Eye Movement with Correction

Mercier, O., Sulai, Y., Mackenzie, K., Zannoli, M., Hillis, J., Nowrouzezahrai, D., & Lanman, D. (2017). Fast gaze-contingent optimal decompositions for multifocal displays. *ACM Transactions on Graphics (TOG)*, *36*(6), 237.

George-Alex Koulieris

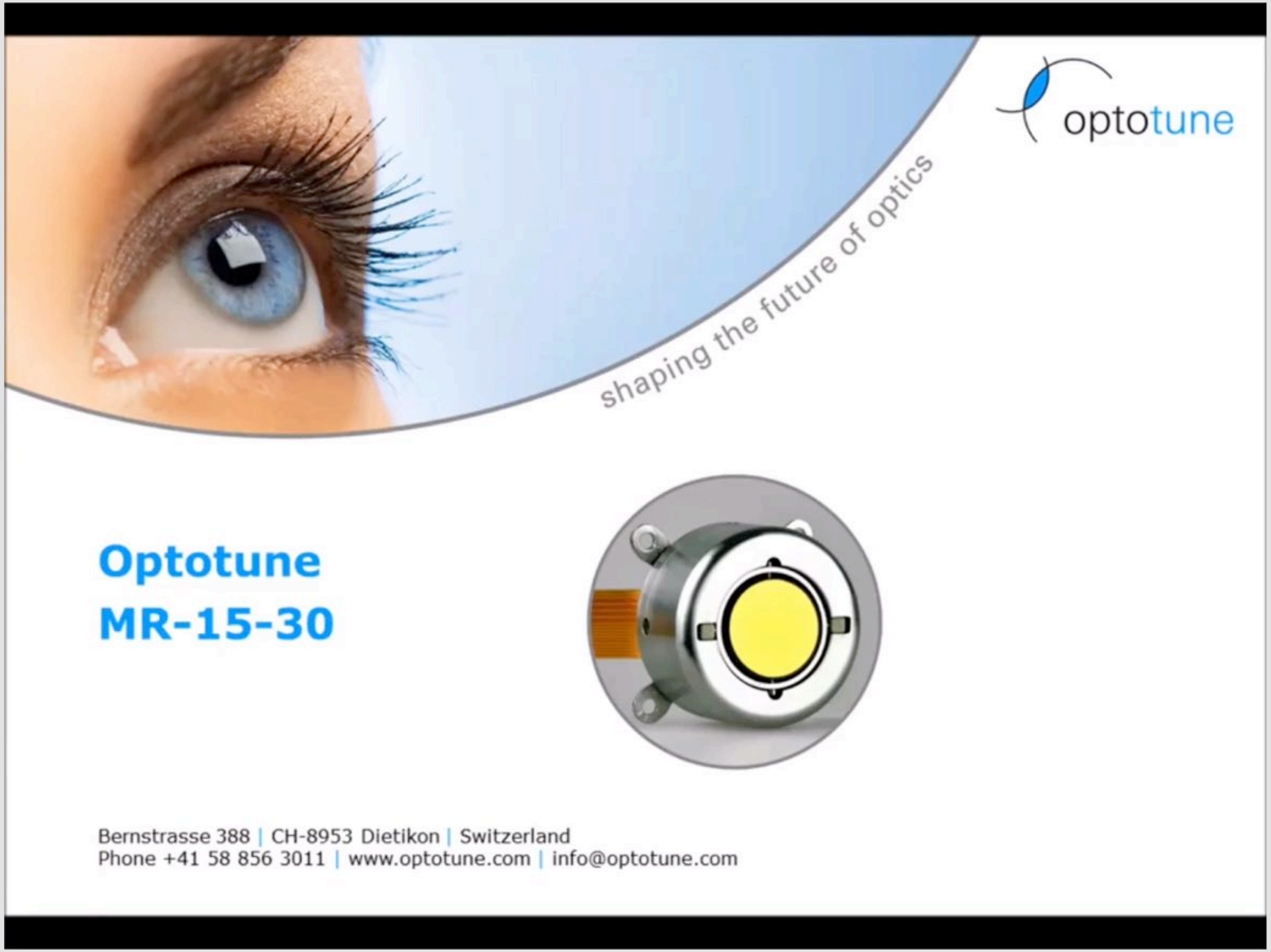
Multifocal scanned voxel displays



McQuaide, S. C., Seibel, E. J., Kelly, J. P., Schowengerdt, B. T., & Furness III, T. A. (2003). A retinal scanning display system that produces multiple focal planes with a deformable membrane mirror. *Displays*, *24*(2), 65-72.

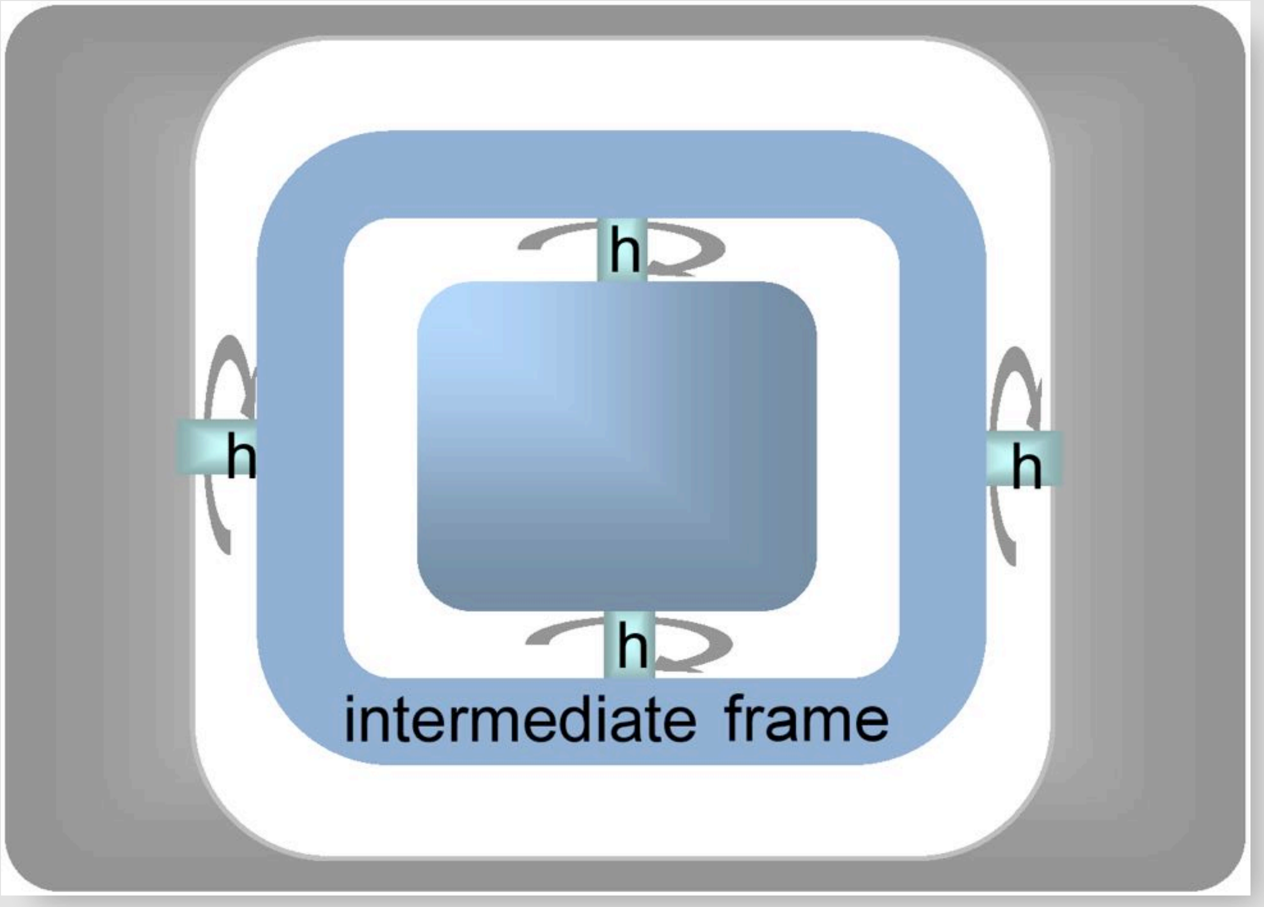
George-Alex Koulieris

Dual axis scanning mirror



Optotune

Principle of operation



Hainich & Bimber, 2017

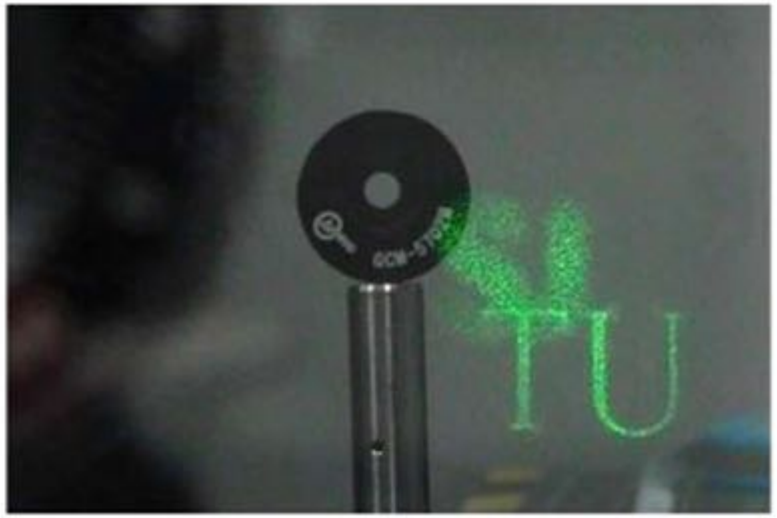
Liquid lenses

Optotune

George-Alex Koulieris

Focusing at different distances

(a)

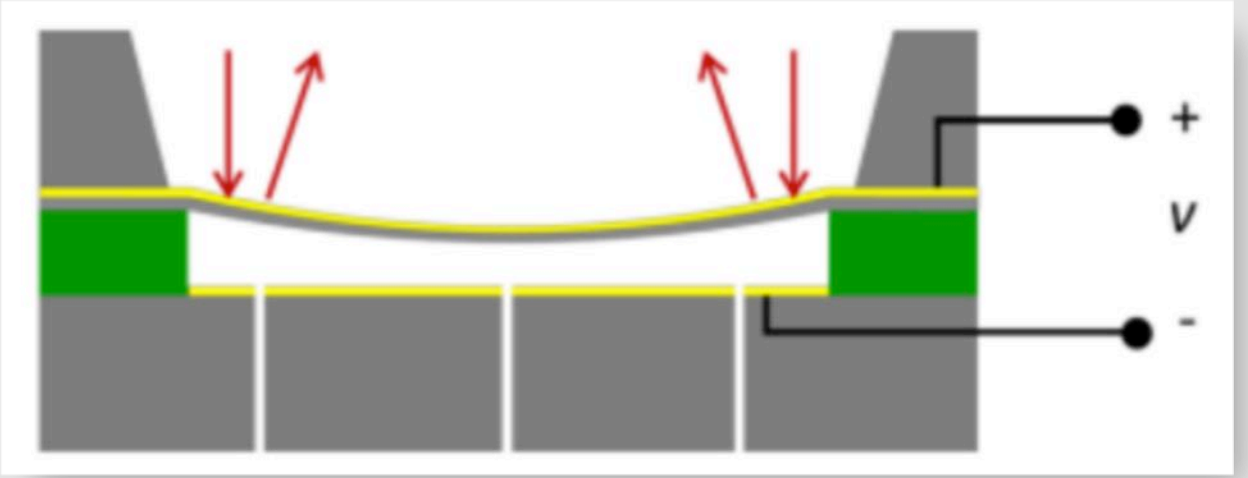


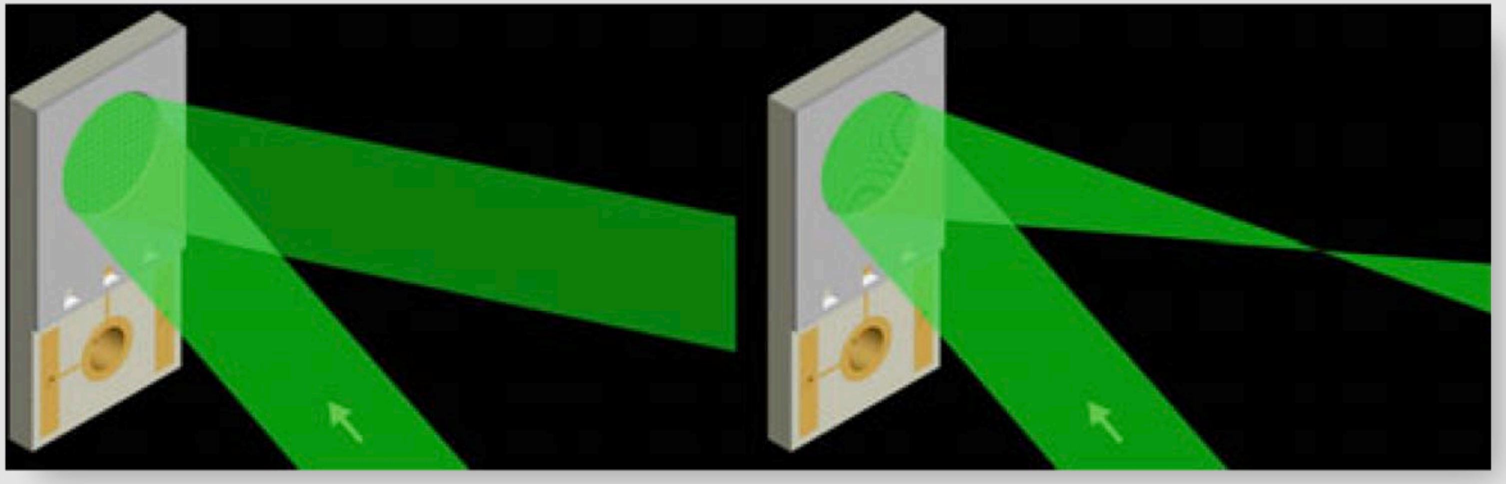
(b)

Xuan Wang

George-Alex Koulieris

Deformable membrane mirrors

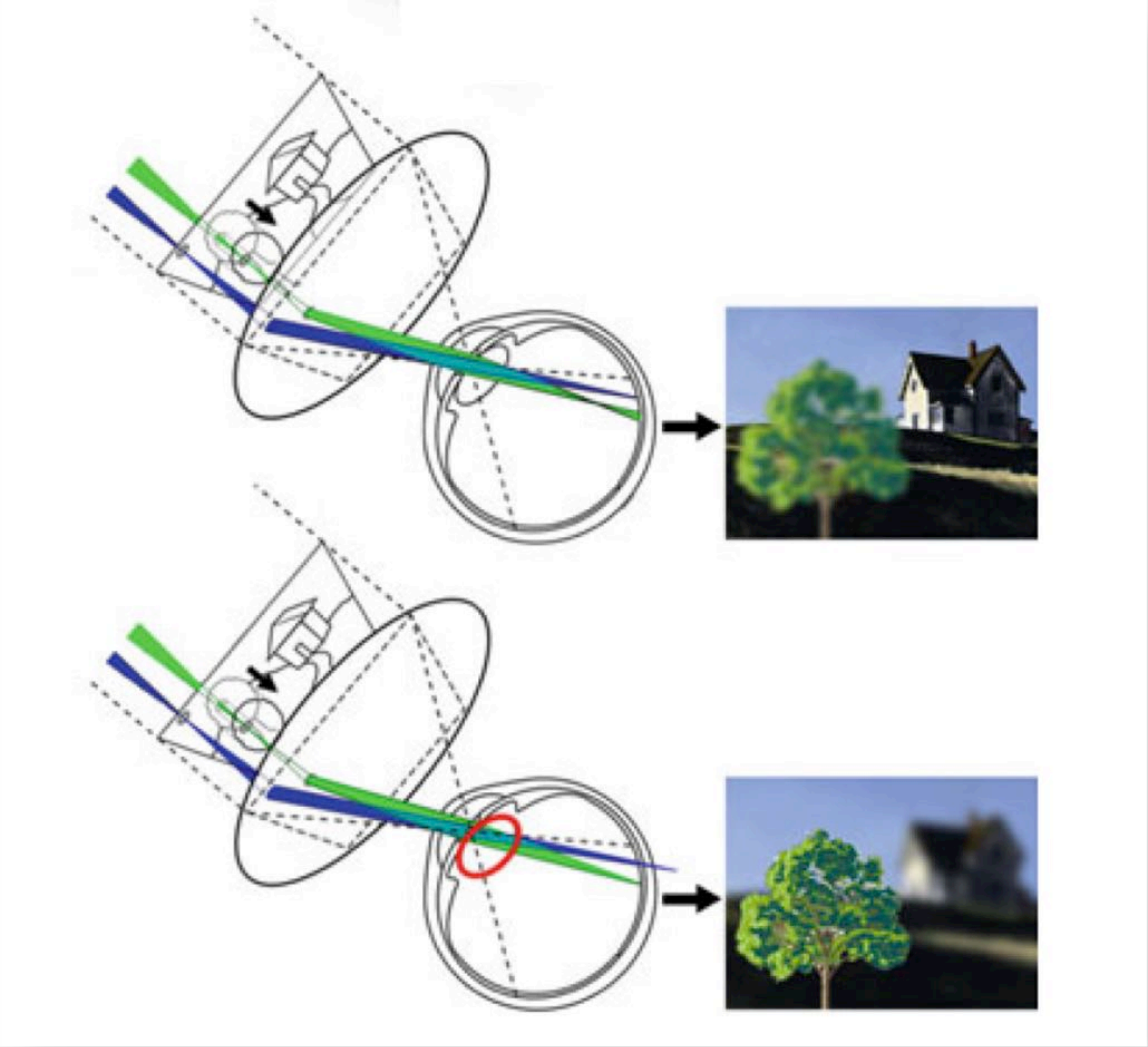




Schowengerdt & Seibel, 2012

McQuaide, S. C., Seibel, E. J., Kelly, J. P., Schowengerdt, B. T., & Furness III, T. A. (2003). A retinal scanning display system that produces multiple focal planes with a deformable membrane mirror. *Displays*, *24*(2), 65-72.

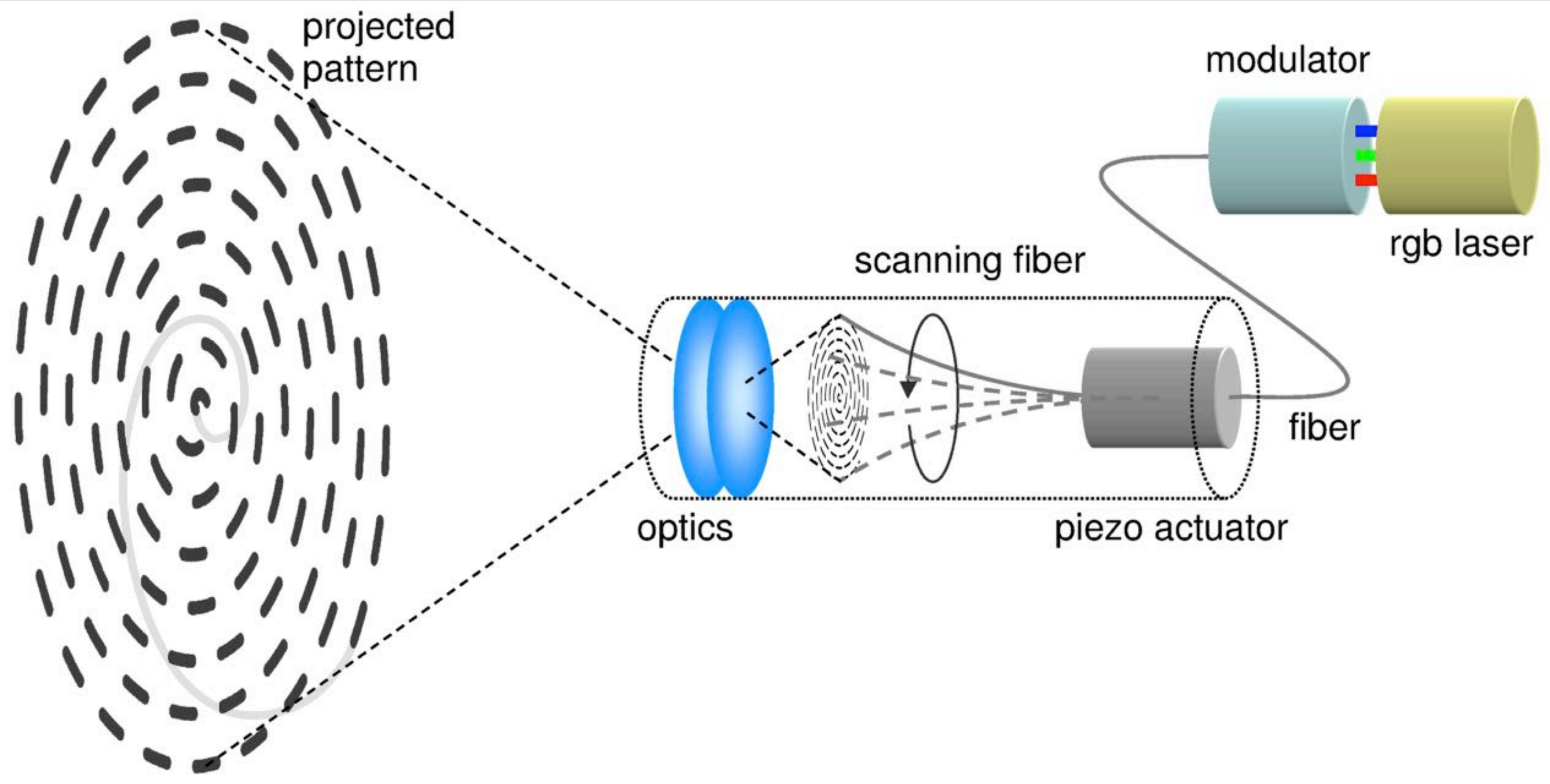
Multifocal scanned voxel displays



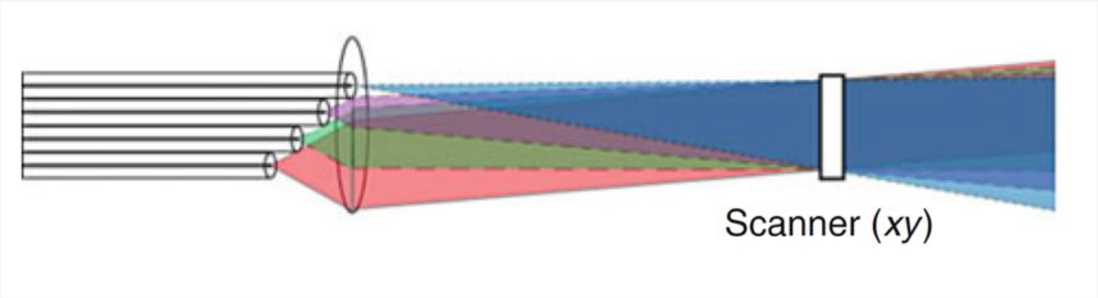
McQuaide, S. C., Seibel, E. J., Kelly, J. P., Schowengerdt, B. T., & Furness III, T. A. (2003). A retinal scanning display system that produces multiple focal planes with a deformable membrane mirror. *Displays*, *24*(2), 65-72.

George-Alex Koulieris

Scanning fiber projector



Hainich & Bimber, 2017



Schowengerdt & Seibel, 2012

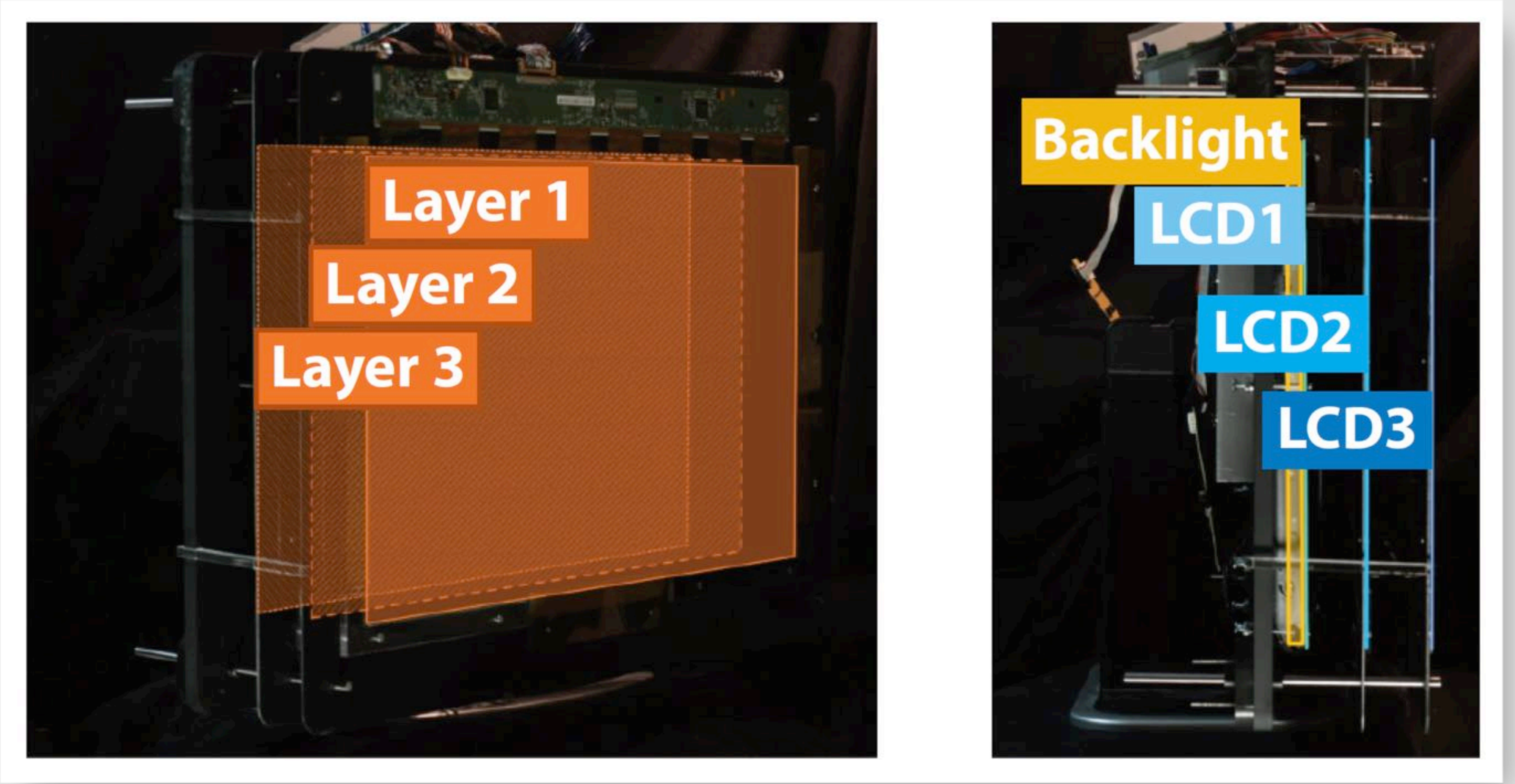
George-Alex Koulieris

Light field displays

- Emit a 4-dimensional distribution of light rays
  - 2D on the display
  - Another 2D horizontal & vertical angle of each pixel
- Each light ray carries radiance at some location into a specific direction

Lanman, D., Hirsch, M., Kim, Y., & Raskar, R. (2010, December). Content-adaptive parallax barriers: optimizing dual-layer 3D displays using low-rank light field factorization. In *ACM Transactions on Graphics (TOG)* (Vol. 29, No. 6, p. 163). ACM.

Light field displays

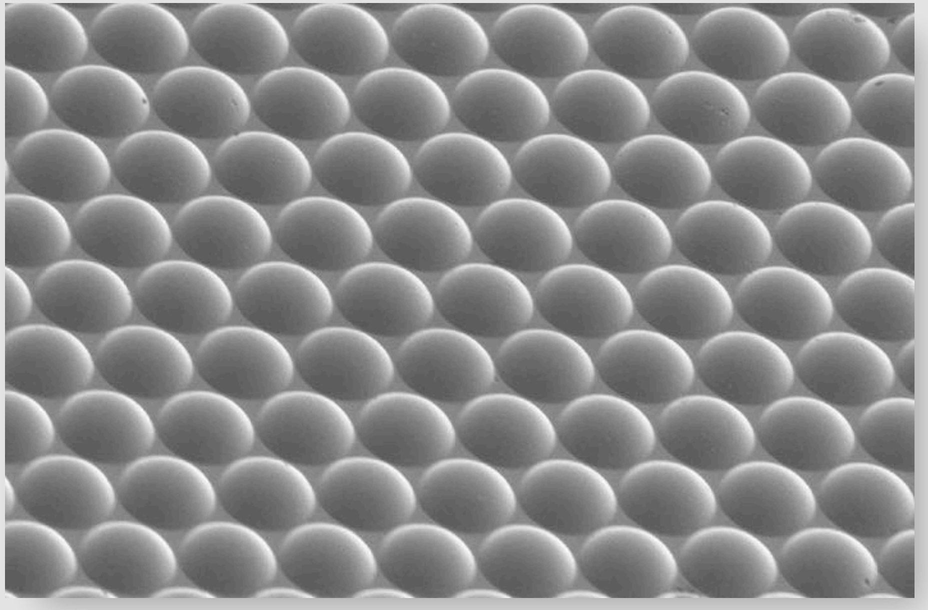


Wetzstein, G., Lanman, D., Hirsch, M., & Raskar, R. (2012). Tensor displays: compressive light field synthesis using multilayer displays with directional backlighting.

George-Alex Koulieris

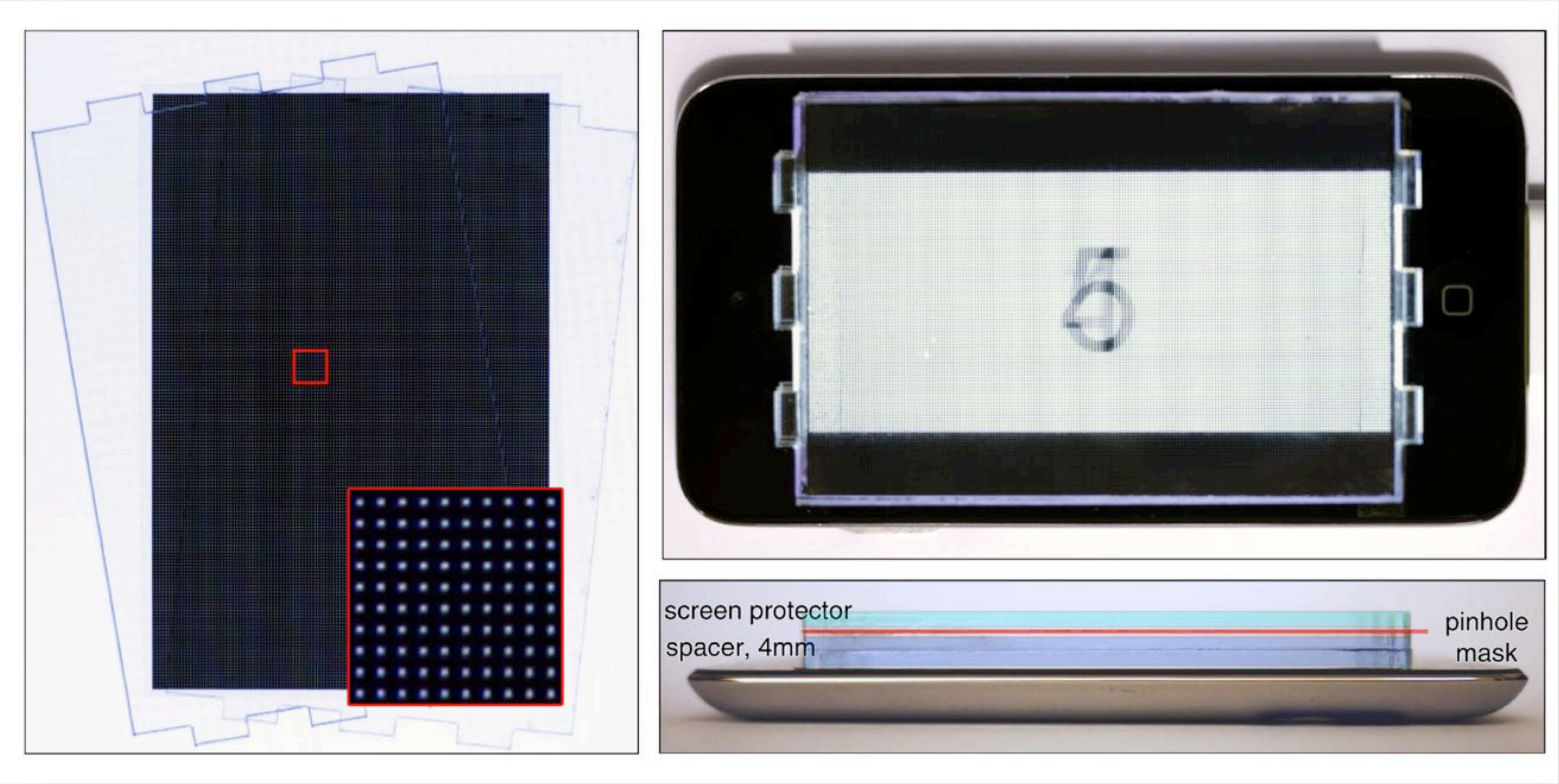
Example construction

- Sandwich a microlens array between an LCD-pair stack
- Perform light beam steering and modulation



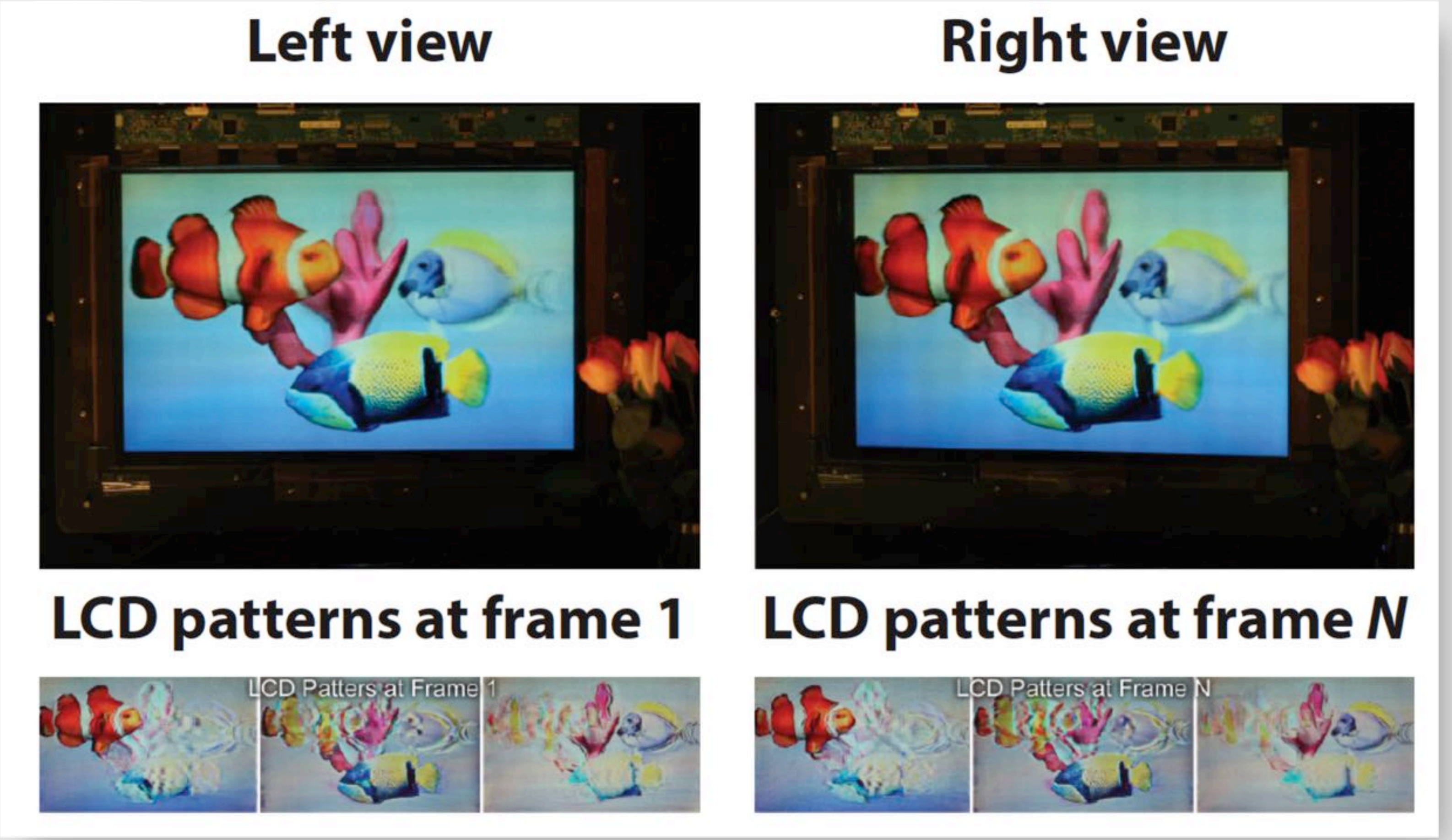
Planettech International

Pinhole parallax barrier
5x5 pixels under each pinhole



Huang, F. C., Wetzstein, G., Barsky, B. A., & Raskar, R. (2014). Eyeglasses-free display: towards correcting visual aberrations with computational light field displays. *ACM Transactions on Graphics (TOG)*, *33*(4), 59.

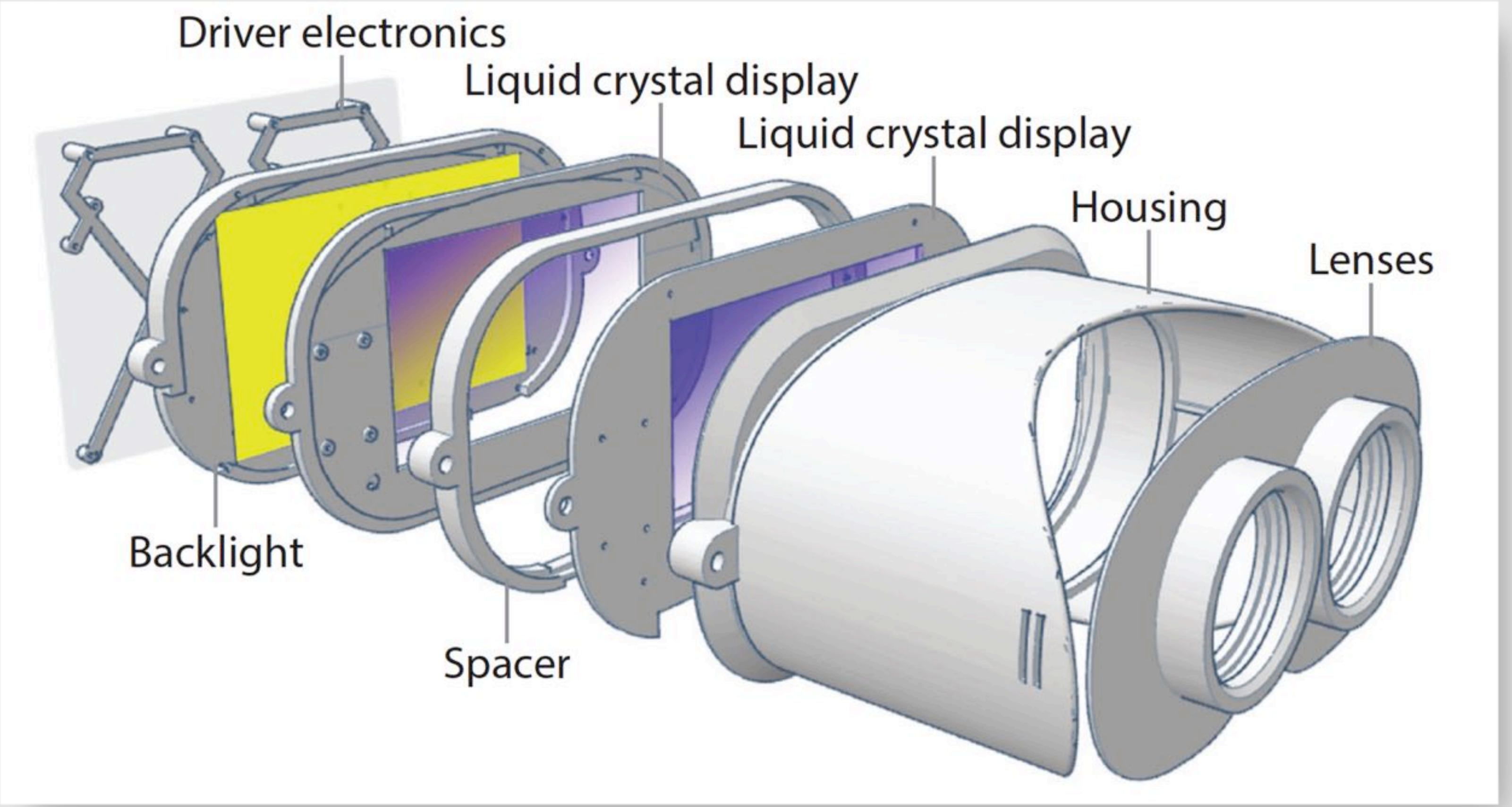
Light field displays



Wetzstein, G., Lanman, D., Hirsch, M., & Raskar, R. (2012). Tensor displays: compressive light field synthesis using multilayer displays with directional backlighting.

George-Alex Koulieris

Wearable light field displays



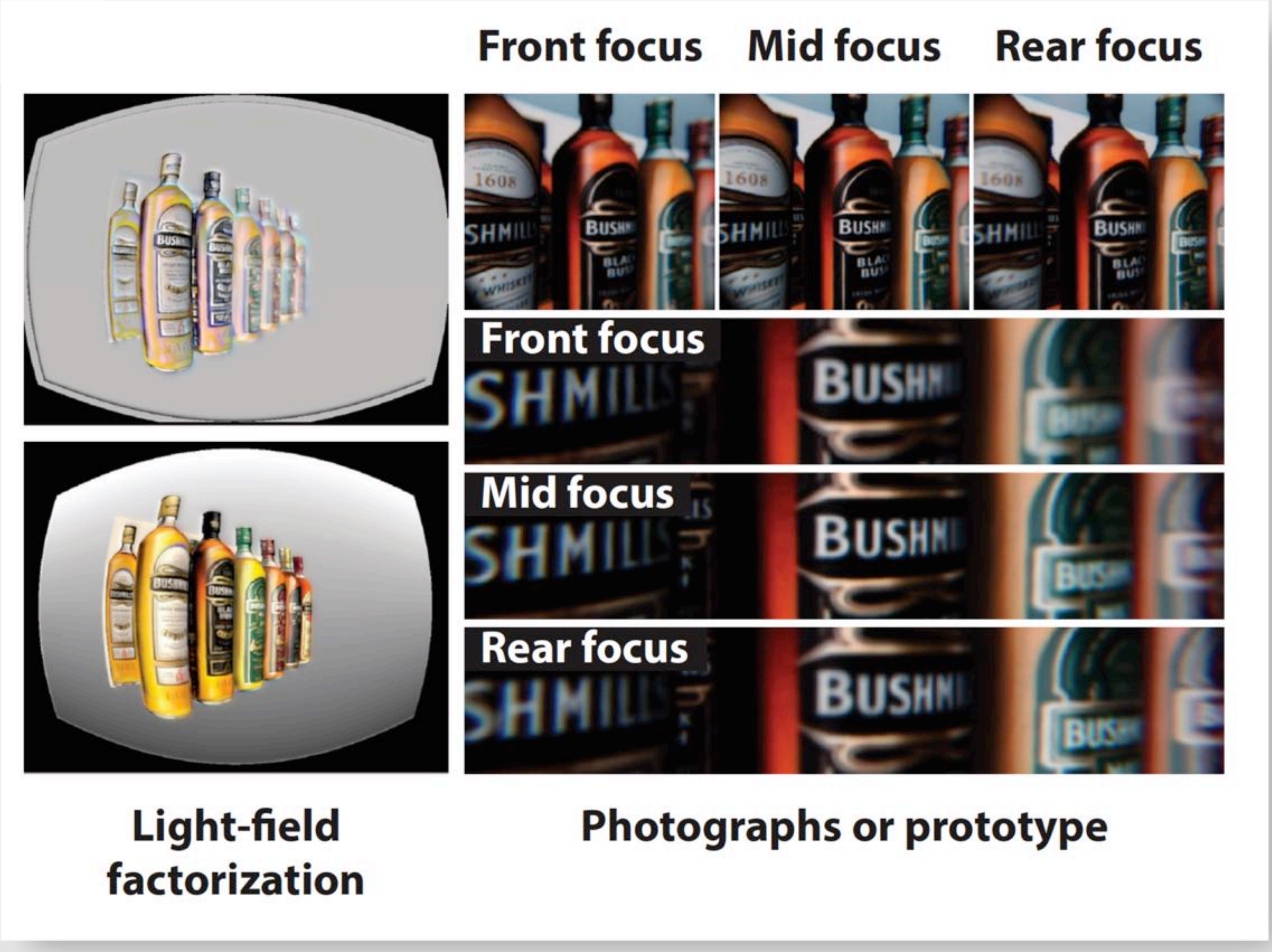
Huang et al., 2015

Lanman, D., & Luebke, D. (2013). Near-eye light field displays. *ACM Transactions on Graphics (TOG)*, *32*(6), 220.

Huang, F. C., Chen, K., & Wetzstein, G. (2015). The light field stereoscope: immersive computer graphics via factored near-eye light field displays with focus cues. *ACM Transactions on Graphics (TOG)*, *34*(4), 60.

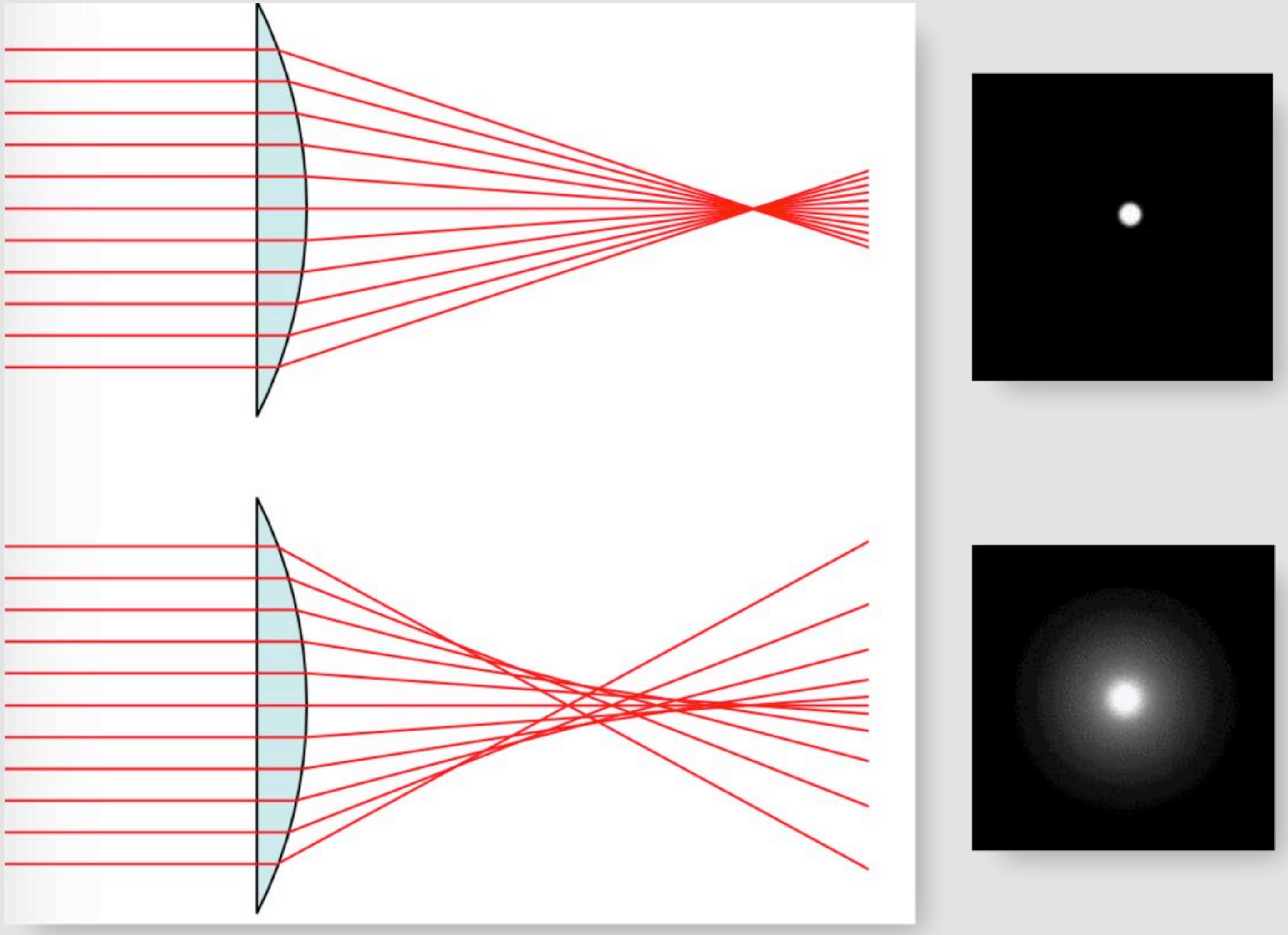
George-Alex Koulieris

Wearable light field displays



George-Alex Koulieris

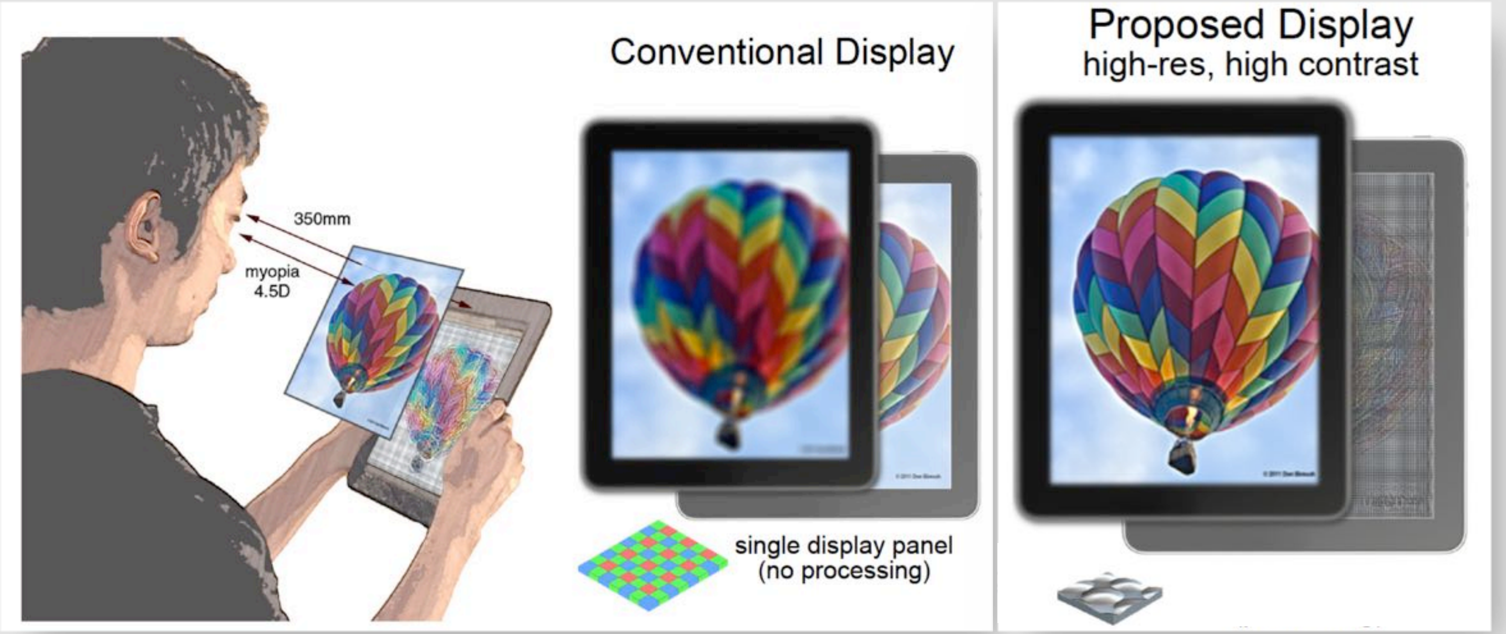
Spherical aberrations



Mglg

George-Alex Koulieris

Pre-correcting aberrations with light field displays



Huang, F. C., Wetzstein, G., Barsky, B. A., & Raskar, R. (2014). Eyeglasses-free display: towards correcting visual aberrations with computational light field displays. *ACM Transactions on Graphics (TOG)*, *33*(4), 59.

George-Alex Koulieris

Holography

- The ultimate 3D image generation technique
- Exact wave-front reconstruction
- Holograms record and play all characteristics of light waves
  - phase, amplitude, wavelength

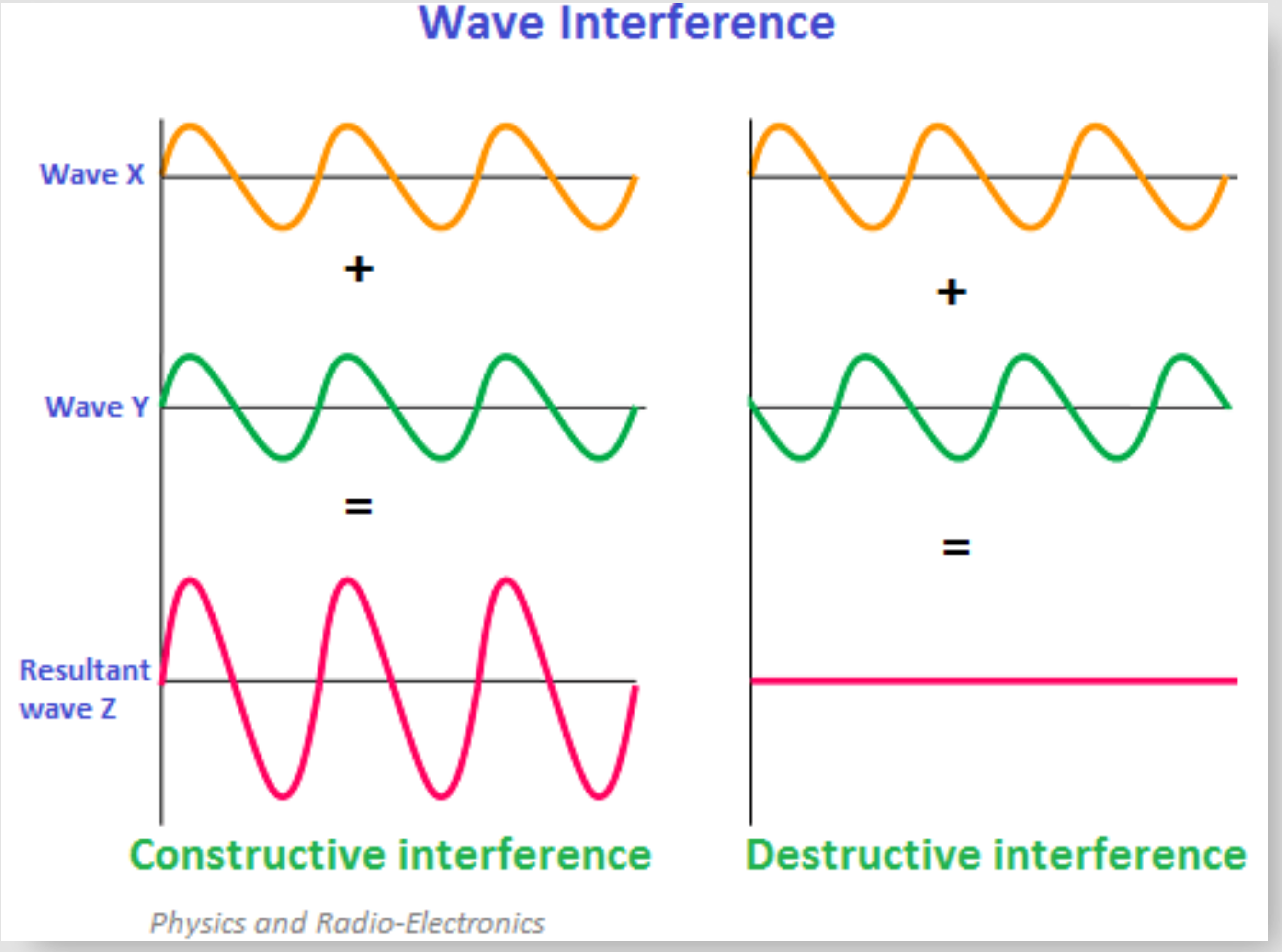
Holography

- Ideally no difference between real object and its hologram
- Recorded using lasers that exhibit coherent monochrome light with regular wave-fronts on photographic plates
- Can use 3 colored lasers for color reproduction
- Computer generated holograms very promising in the -far- future

George-Alex Koulieris

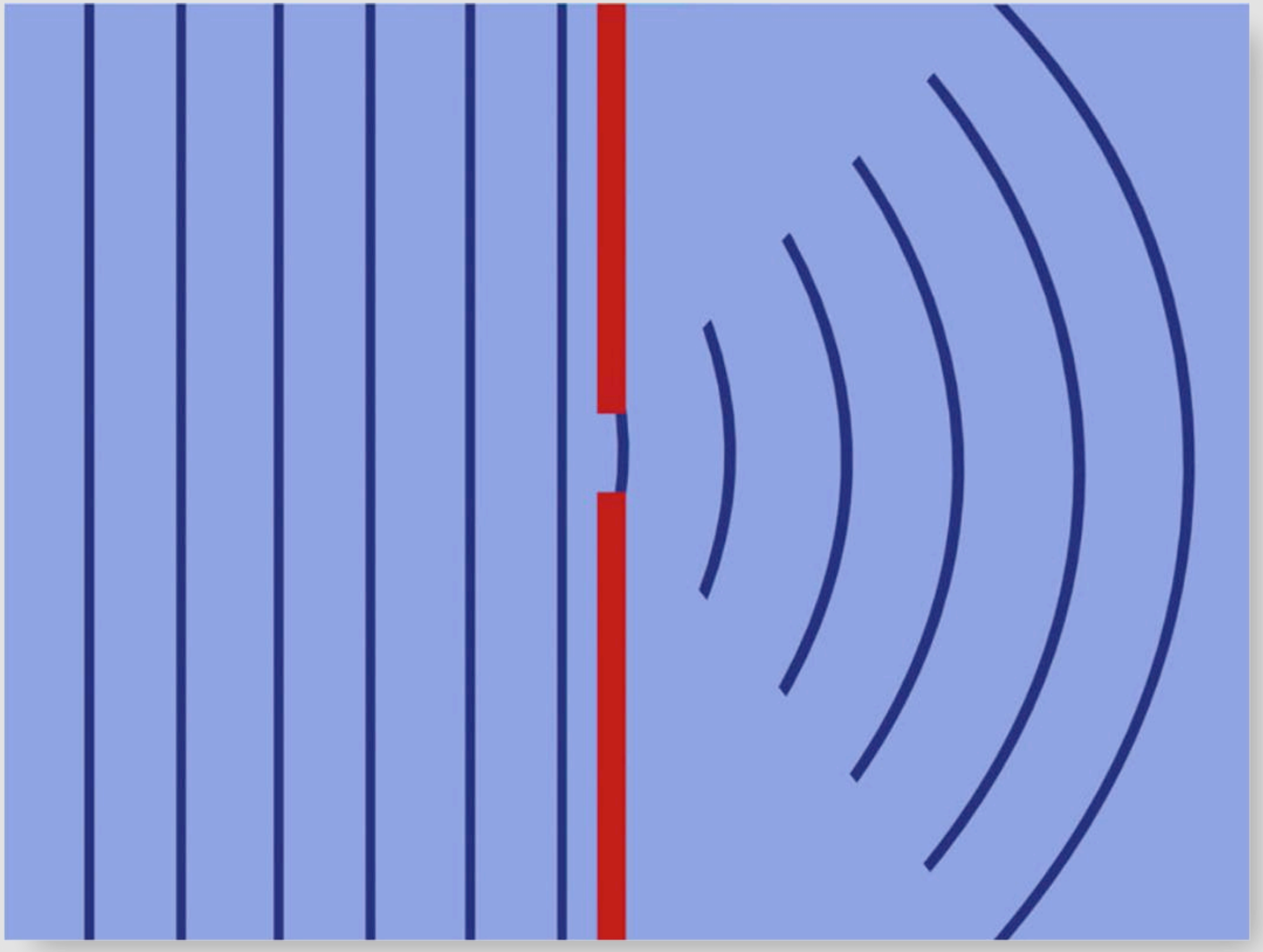
Principles

- Interference



Principles

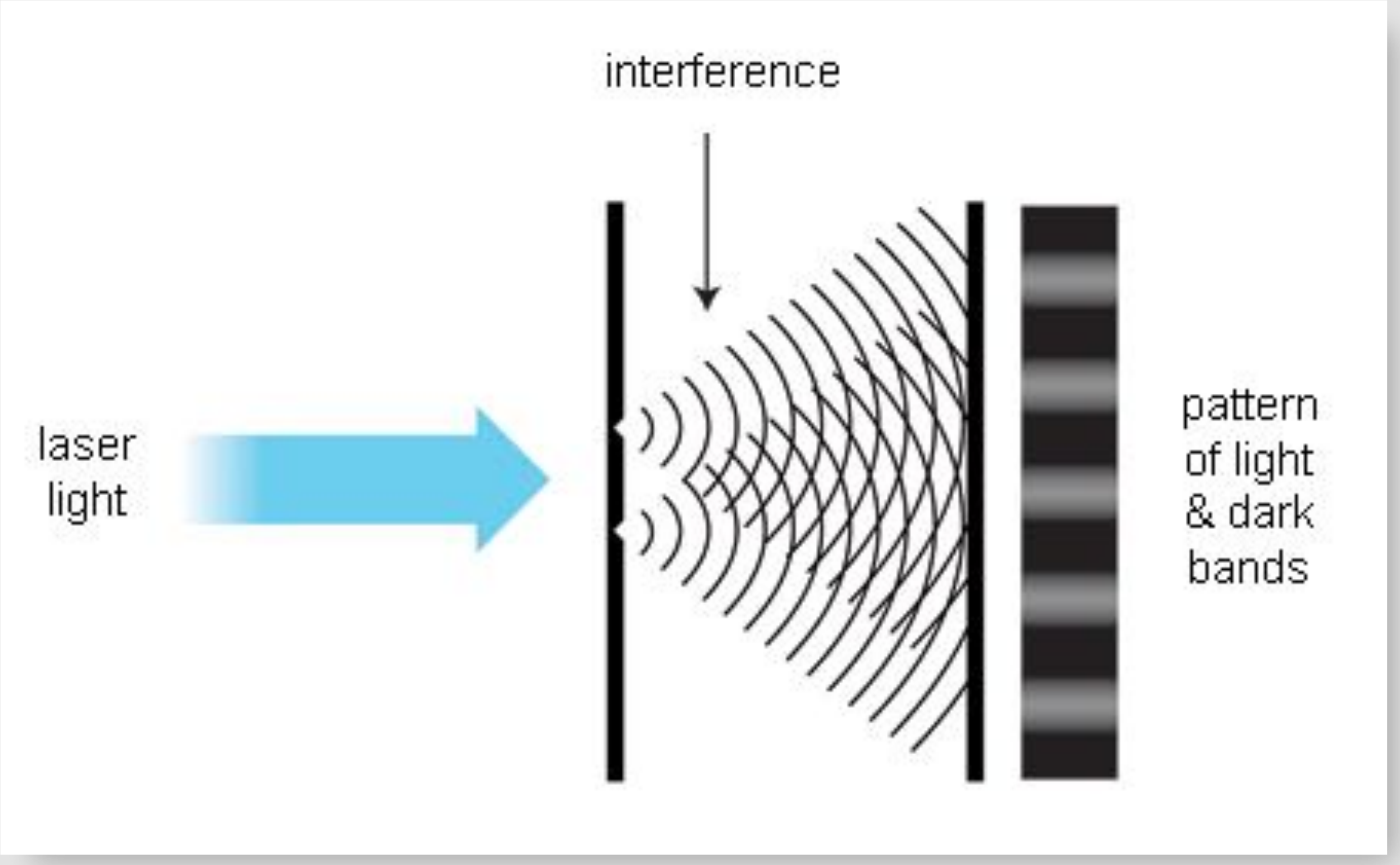
- Interference
- Diffraction



George-Alex Koulieris

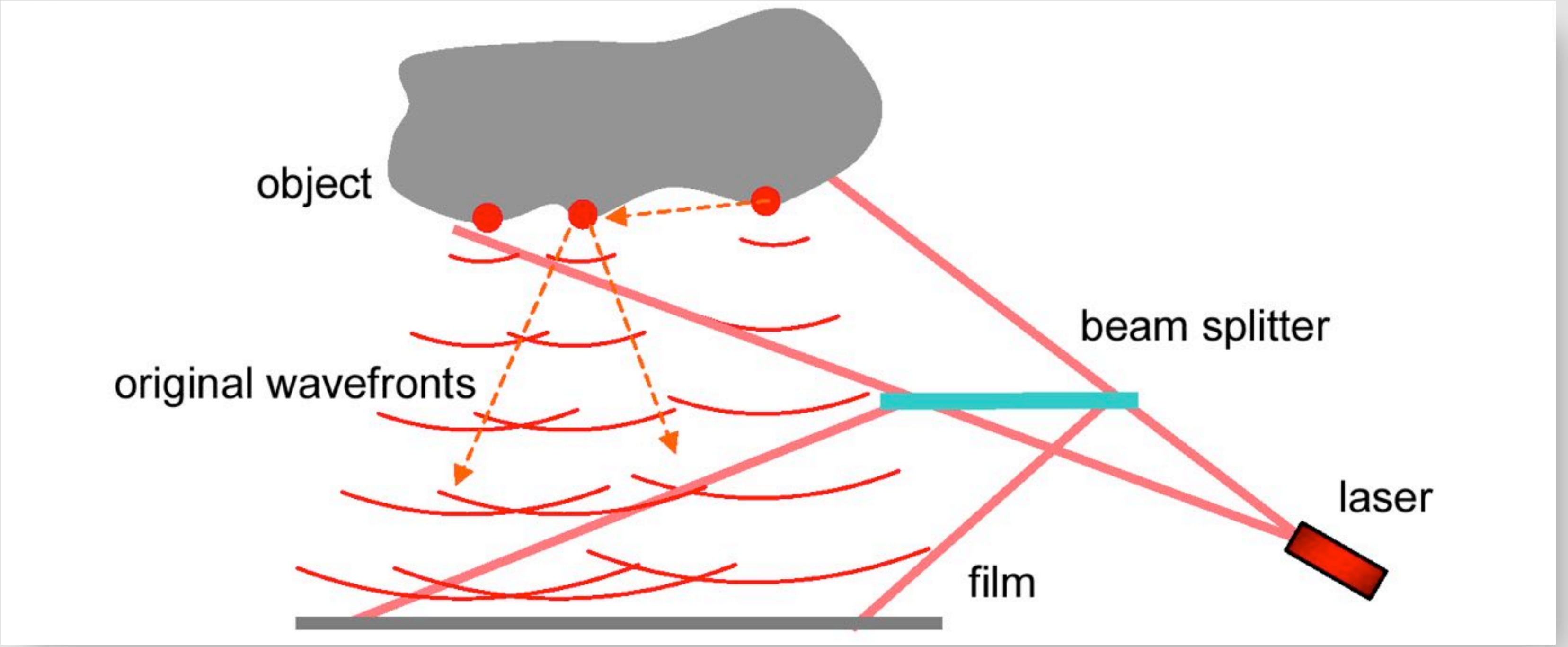
Principles

- Interference
- Diffraction
- Fringe pattern superposition



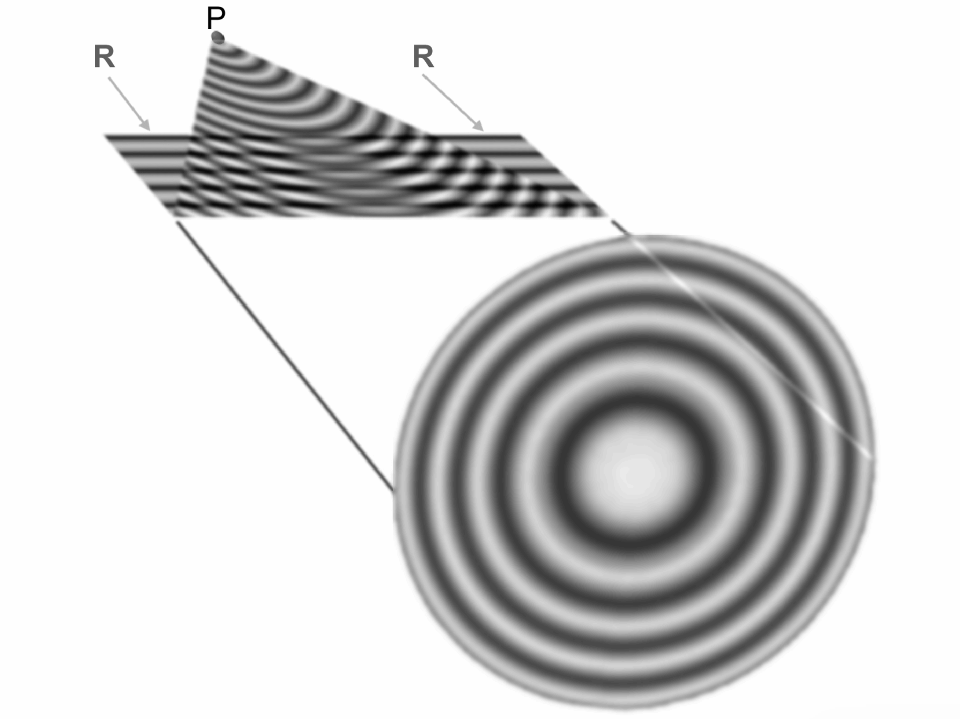
George-Alex Koulieris

Recording holograms

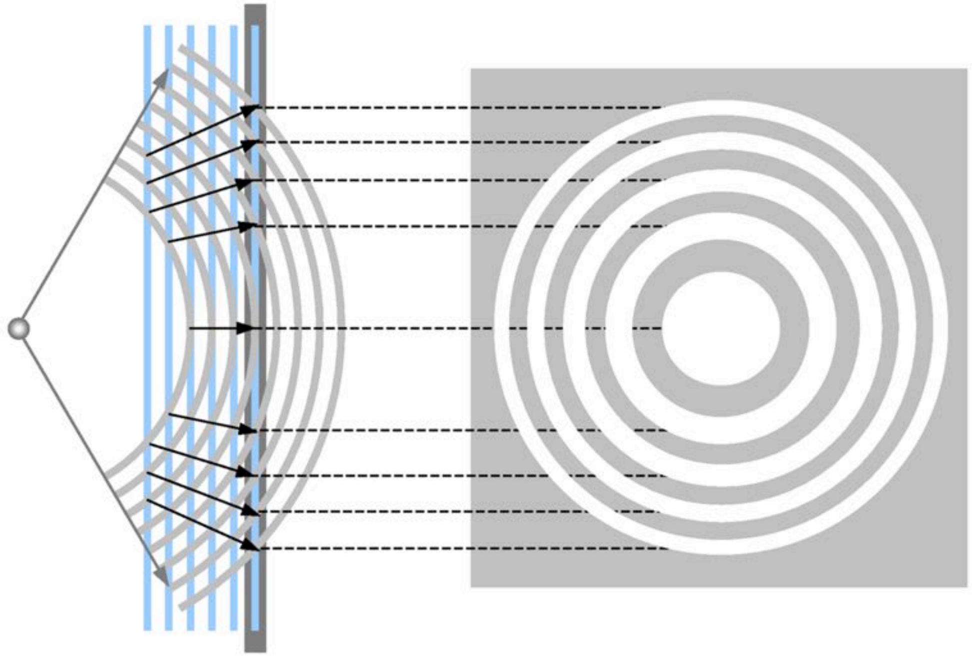


Hainich & Bimber, 2017

Recorded fringe patterns



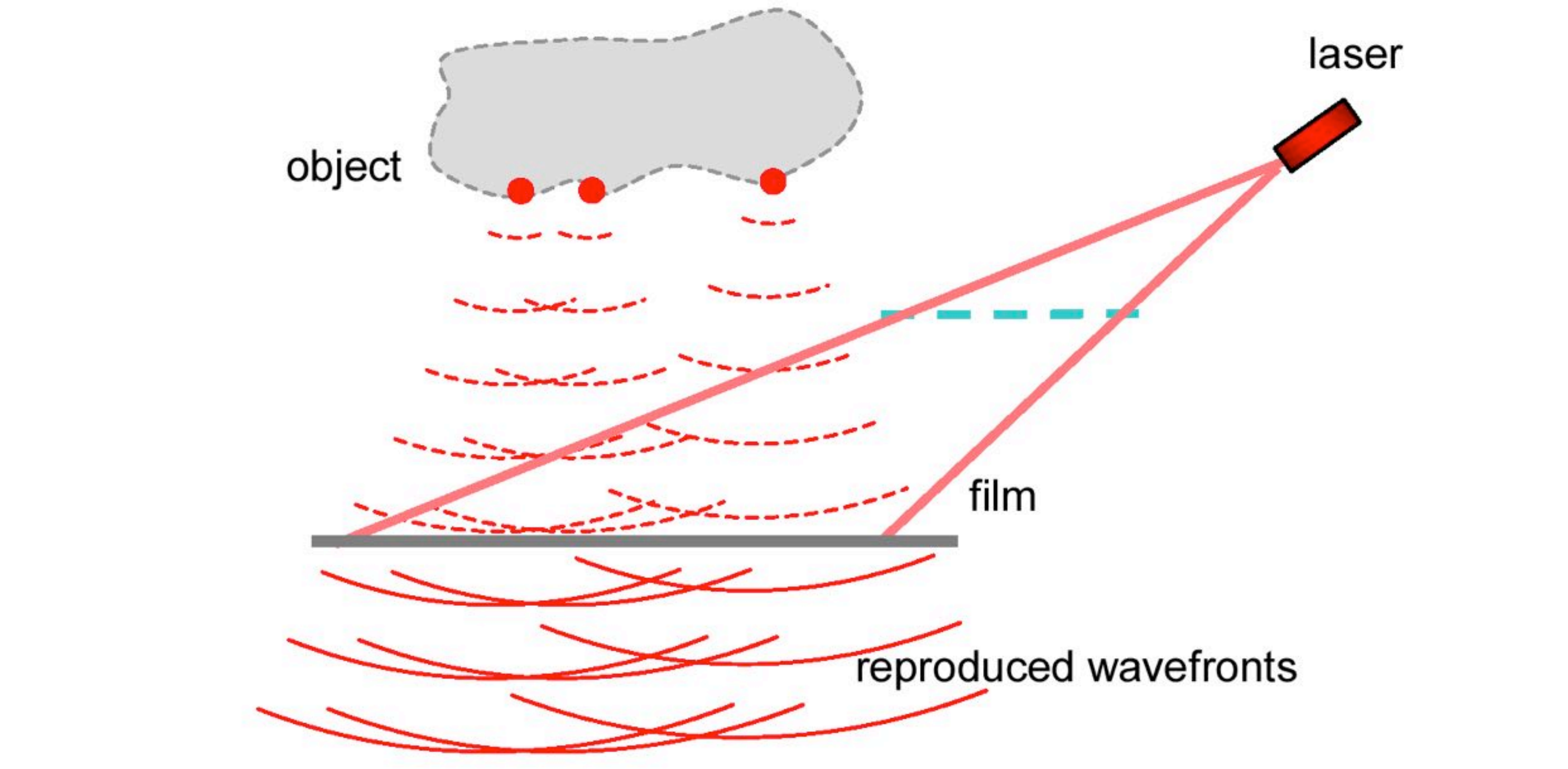
Hainich & Bimber, 2017



Hainich & Bimber, 2017

George-Alex Koulieris

Playing-back holograms

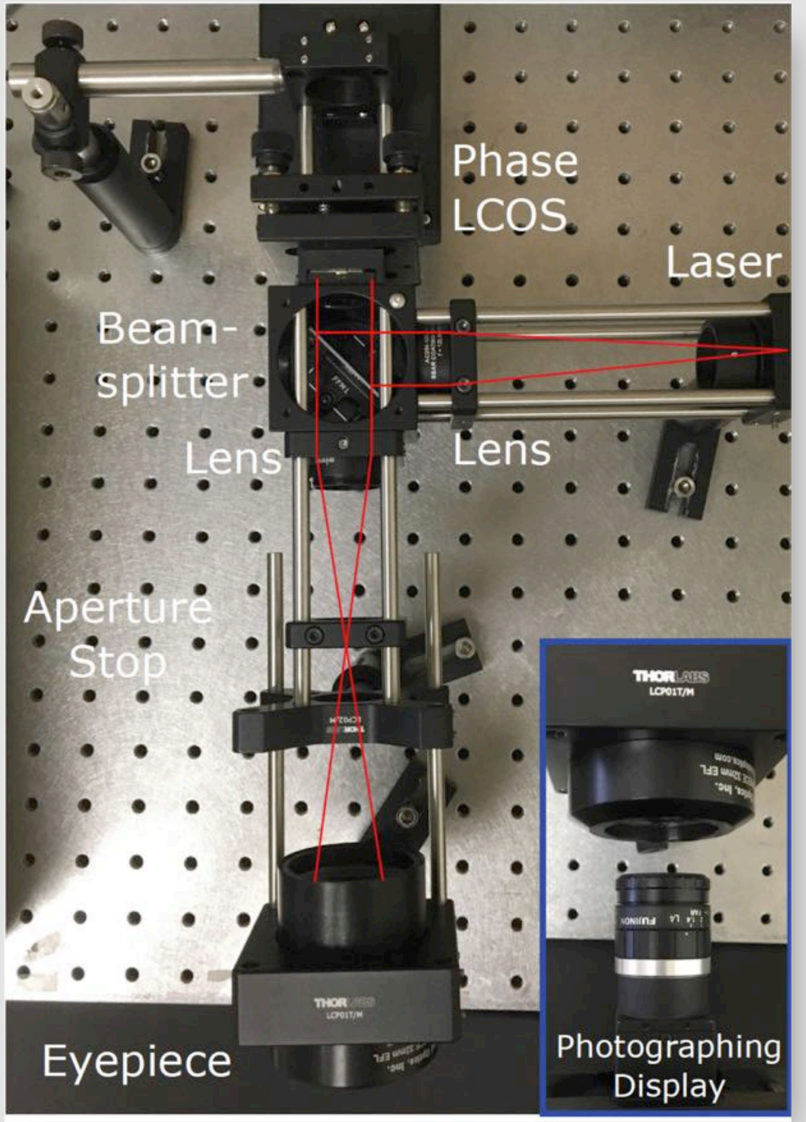


Hainich & Bimber, 2017

Computer generated holograms

- Computer generated fringe patterns
- Use Spatial Light Modulators (SLMs) for display
- DMDs and F-LCDs often used

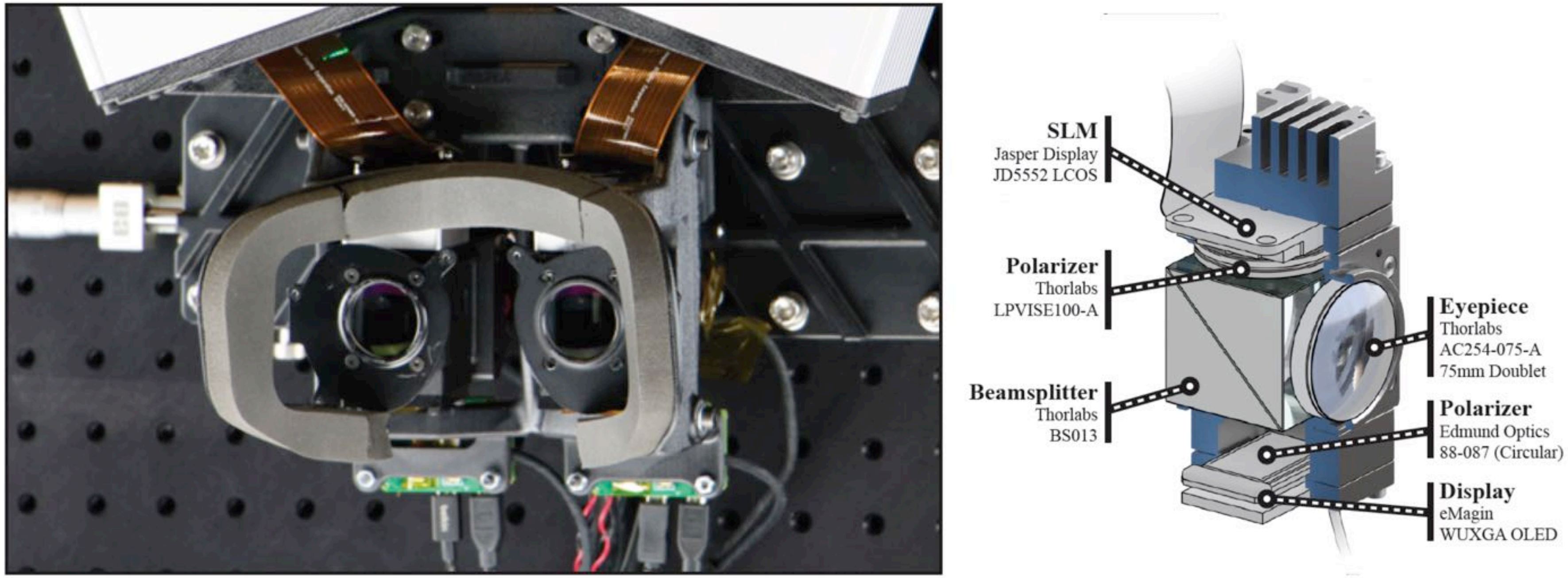
Holographic near-eye displays



Maimone, A., Georgiou, A., & Kollin, J. S. (2017). Holographic near-eye displays for virtual and augmented reality. *ACM Transactions on Graphics (TOG)*, *36*(4), 85.

George-Alex Koulieris

Focal surface displays

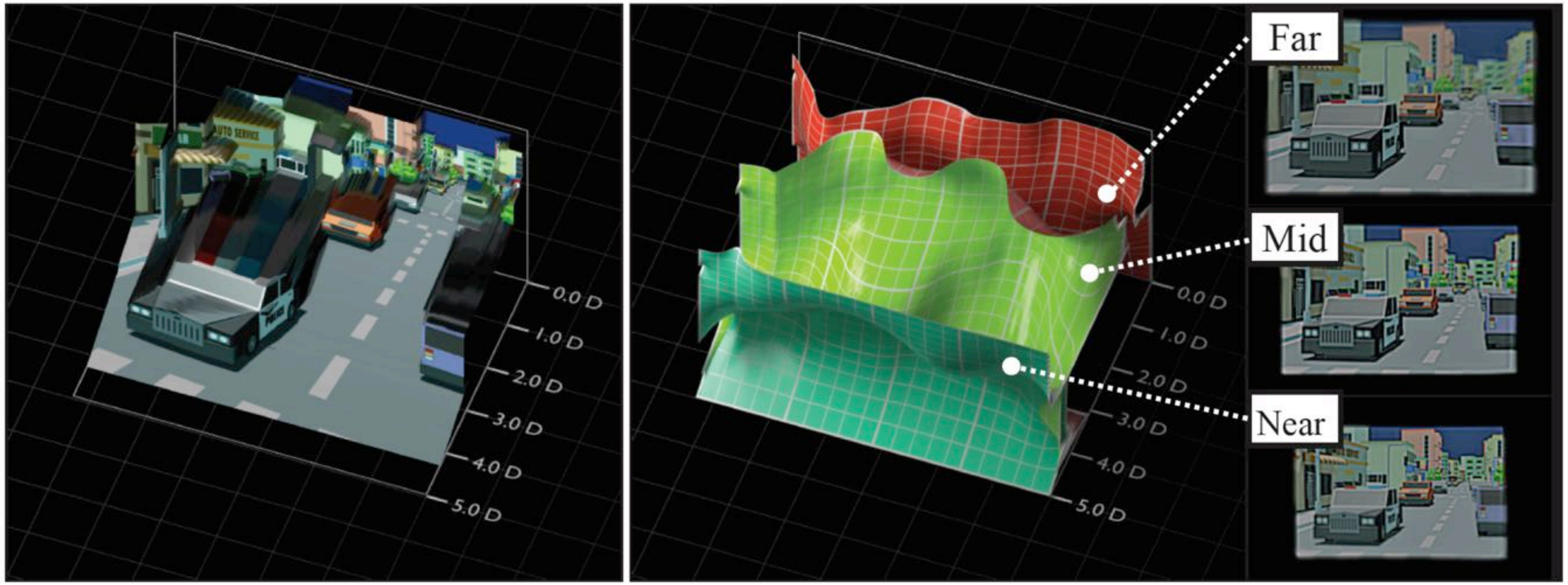


(a) Construction of the Prototype **(b)** Arrangement of the Optical Components

Matsuda, N., Fix, A., & Lanman, D. (2017). Focal surface displays. *ACM Transactions on Graphics (TOG)*, *36*(4), 86.

George-Alex Koulieris

Focal surface displays

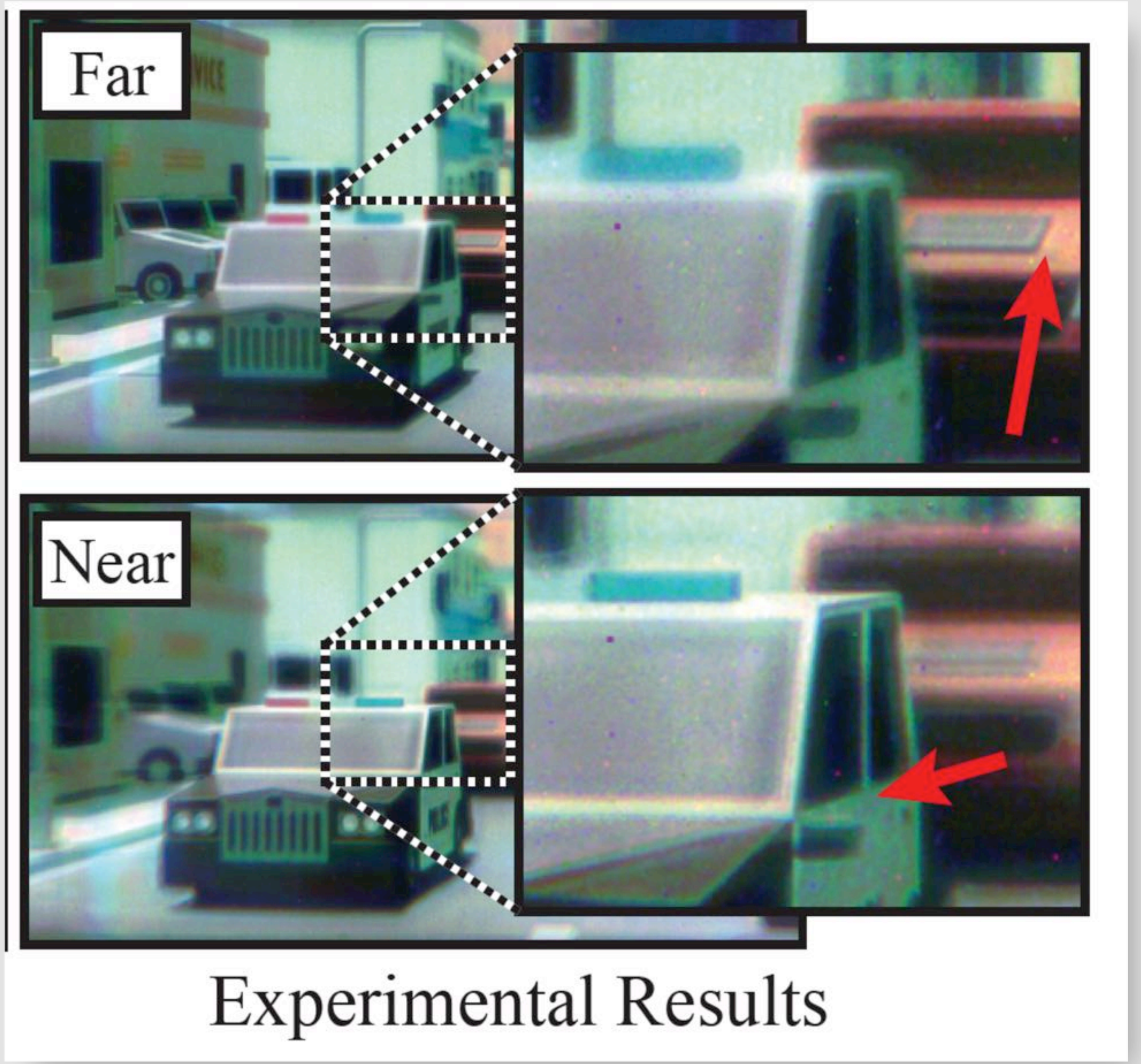


Target Scene
Focal Surface and Color Decomposition

Matsuda, N., Fix, A., & Lanman, D. (2017). Focal surface displays. *ACM Transactions on Graphics (TOG)*, *36*(4), 86.

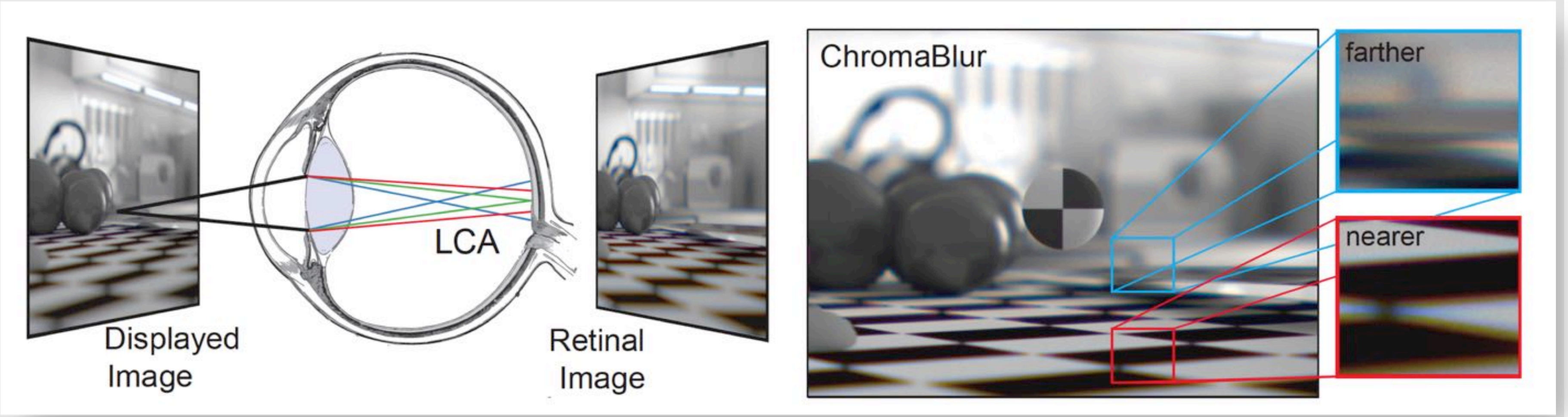
George-Alex Koulieris

Focal surface displays



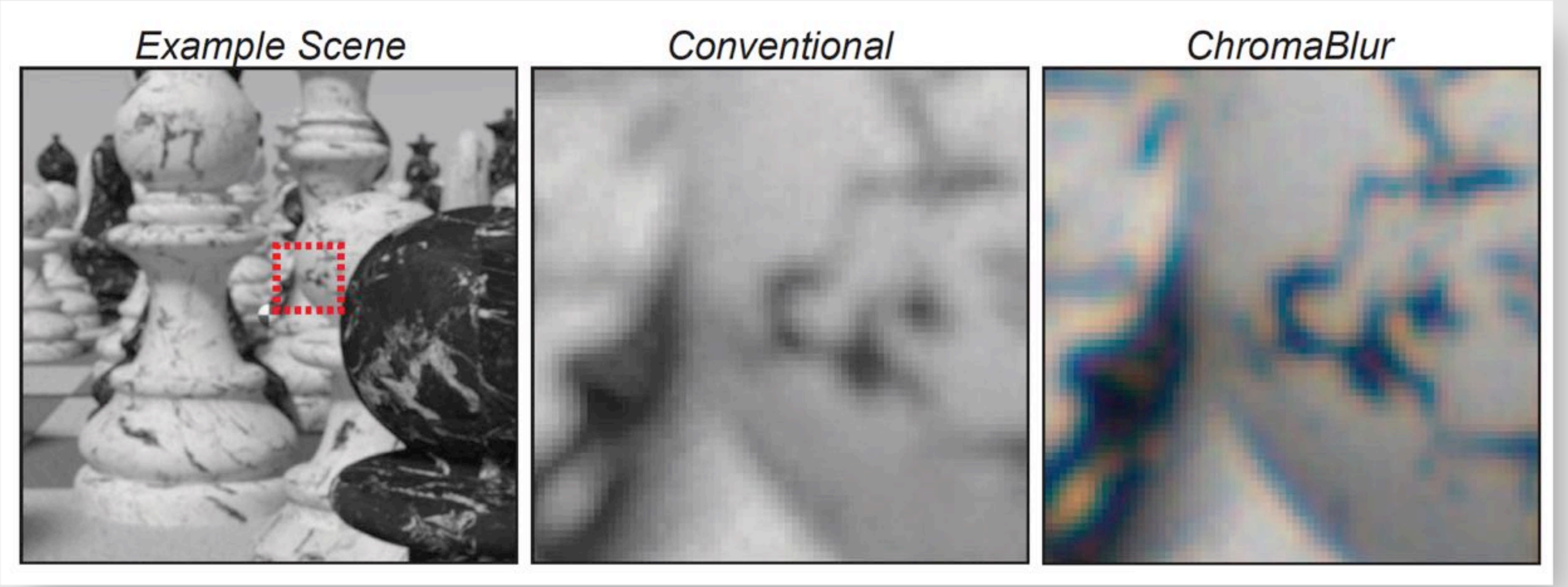
Matsuda, N., Fix, A., & Lanman, D. (2017). Focal surface displays. *ACM Transactions on Graphics (TOG)*, *36*(4), 86.

Rendering chromatic aberration



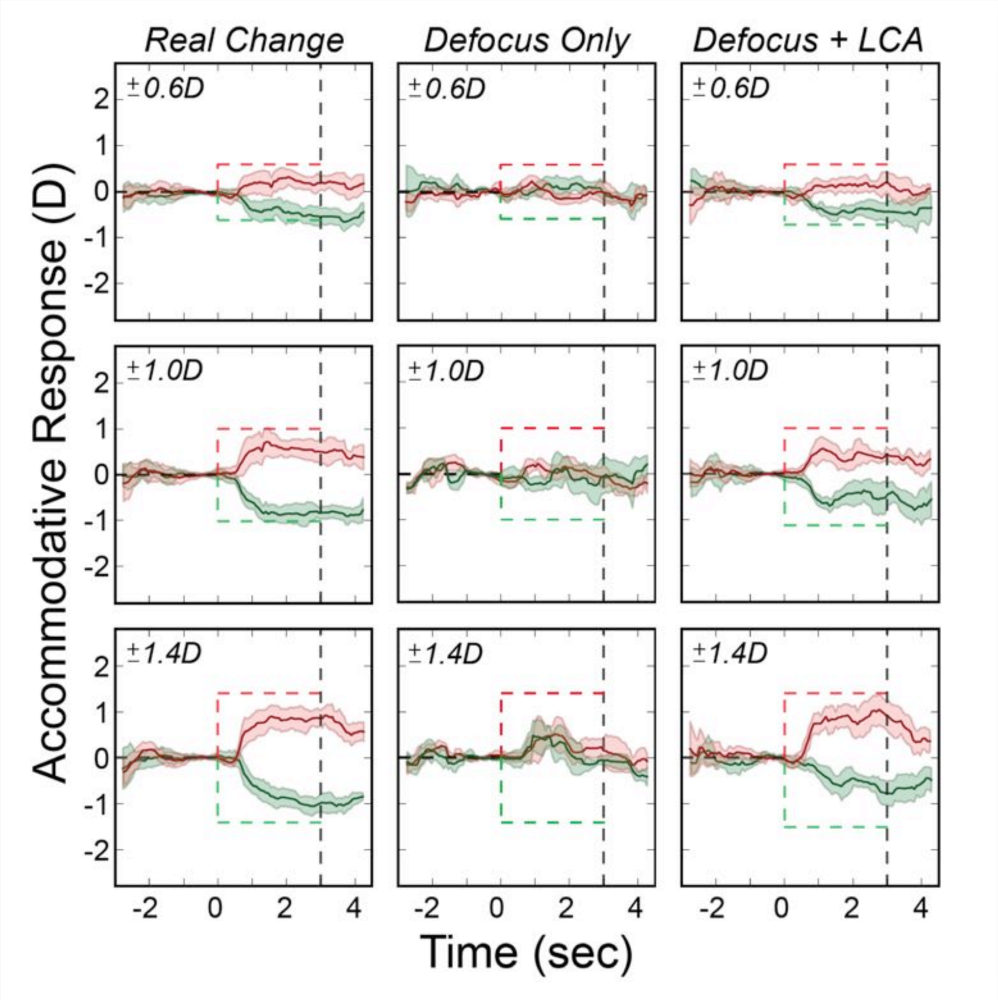
Cholewiak, S. A., Love, G. D., Srinivasan, P. P., Ng, R., & Banks, M. S. (2017). ChromaBlur: rendering chromatic eye aberration improves accommodation and realism. *ACM transactions on graphics.*, *36*(6), 210.

Rendering chromatic aberration



Cholewiak, S. A., Love, G. D., Srinivasan, P. P., Ng, R., & Banks, M. S. (2017). ChromaBlur: rendering chromatic eye aberration improves accommodation and realism. *ACM transactions on graphics.*, *36*(6), 210.

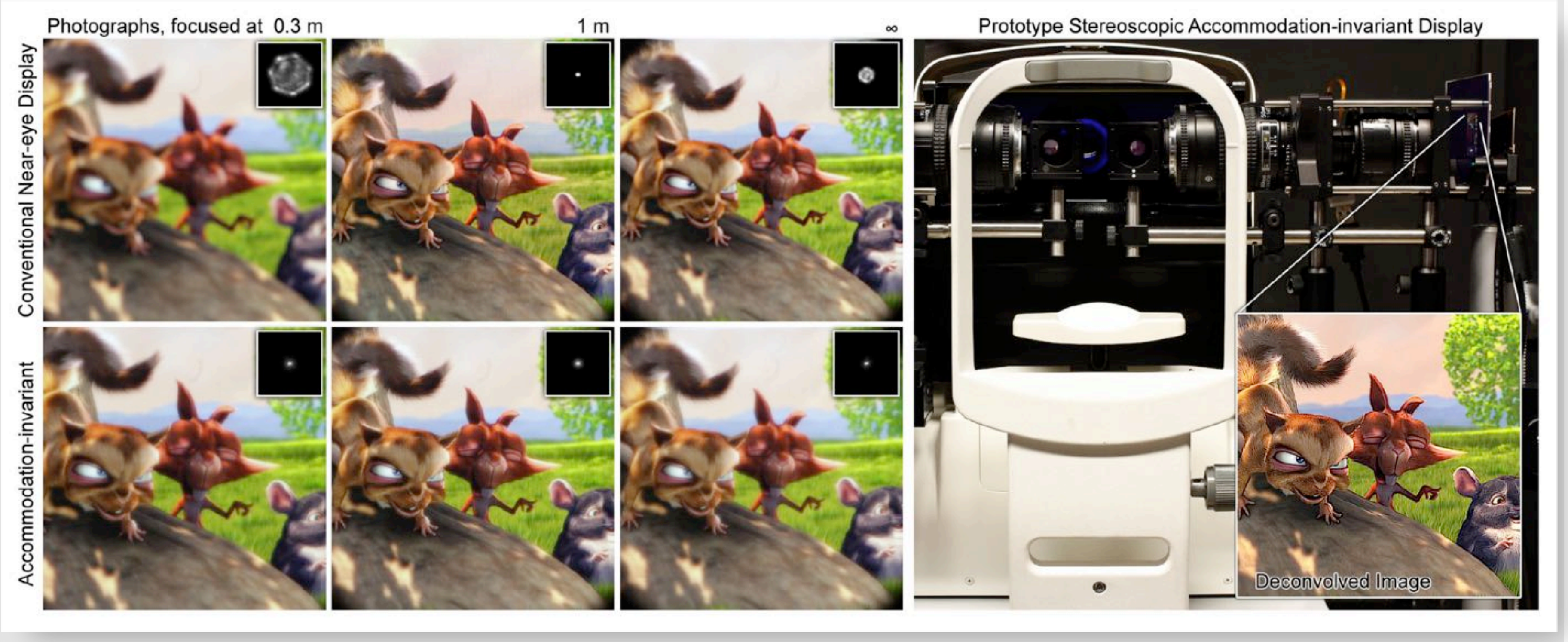
Rendering chromatic aberration



Cholewiak, S. A., Love, G. D., Srinivasan, P. P., Ng, R., & Banks, M. S. (2017). ChromaBlur: rendering chromatic eye aberration improves accommodation and realism. *ACM transactions on graphics.*, *36*(6), 210.

George-Alex Koulieris

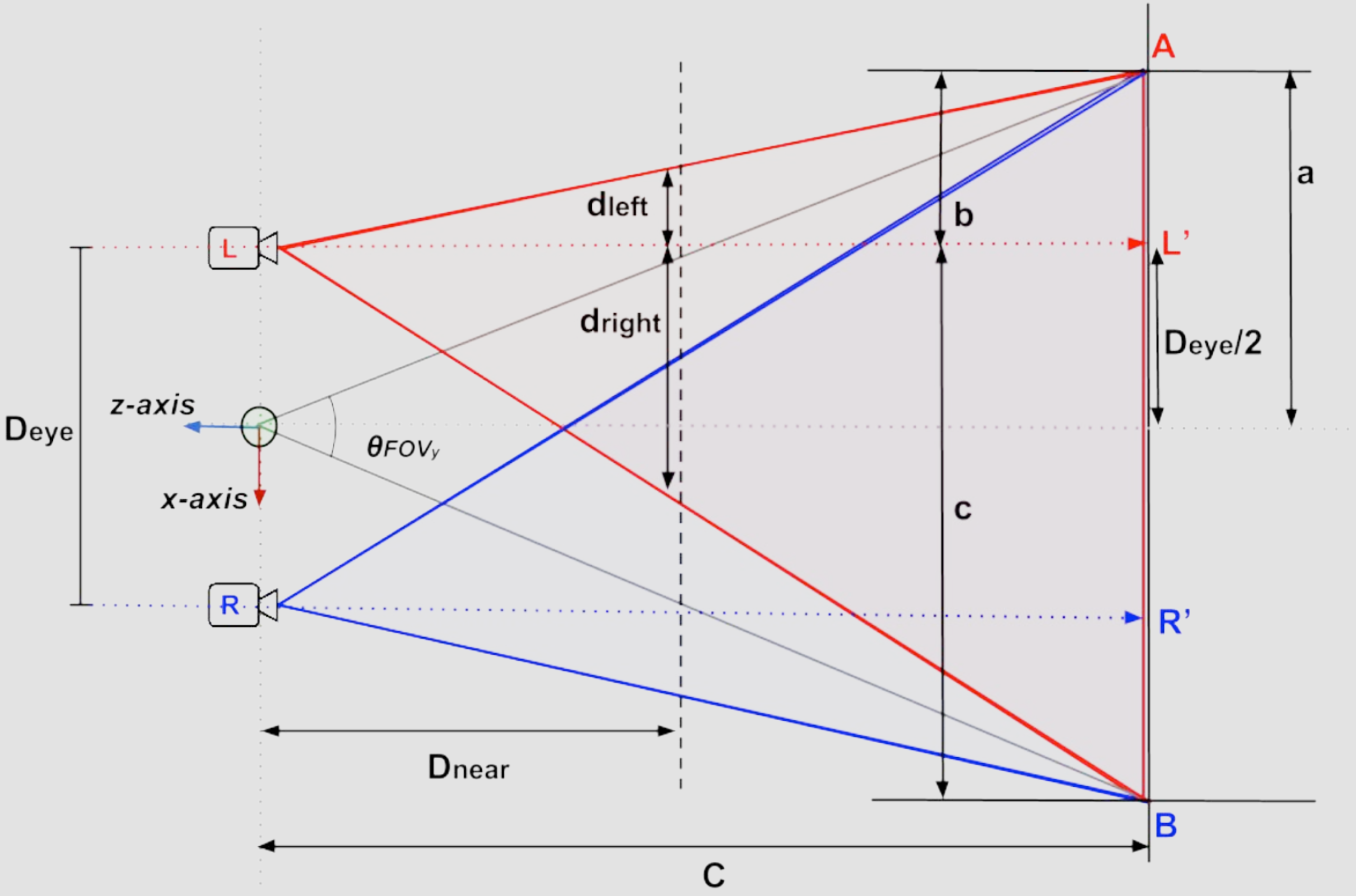
Accommodation invariant displays



Konrad, R., Padmanaban, N., Molner, K., Cooper, E. A., & Wetzstein, G. (2017). Accommodation-invariant computational near-eye displays. *ACM Transactions on Graphics (TOG)*, *36*(4), 88.

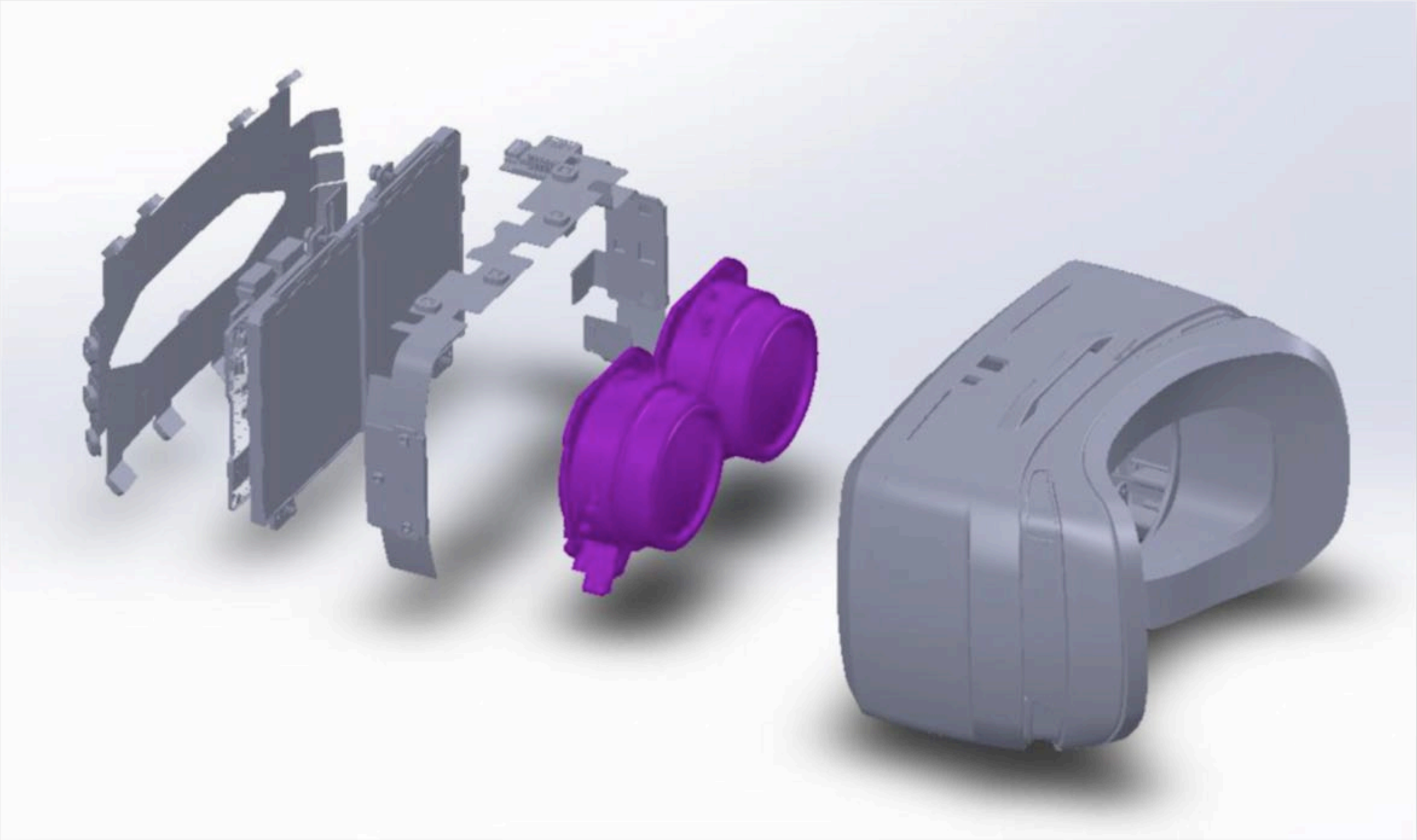
George-Alex Koulieris

Software-only methods



Koulieris, G. A., Drettakis, G., Cunningham, D., & Mania, K. (2016, March). Gaze prediction using machine learning for dynamic stereo manipulation in games. In *Virtual Reality (VR), 2016 IEEE* (pp. 113-120). IEEE.

Lemnis Technologies



George-Alex Koulieris

Thank you

https://vrdisplays.github.io/sigasia2018/

georgios.a.koulieris@durham.ac.uk

Near-Eye

VR/AR Display Technologies

Kaan Akşit

Dec 2018

SIGGRAPH ASIA 2018 TOKYO

NVIDIA

Today


HTC Vive (2016)


Daqri (2017)


Google Cardboard (2016)


Microsoft Hololens (2017)

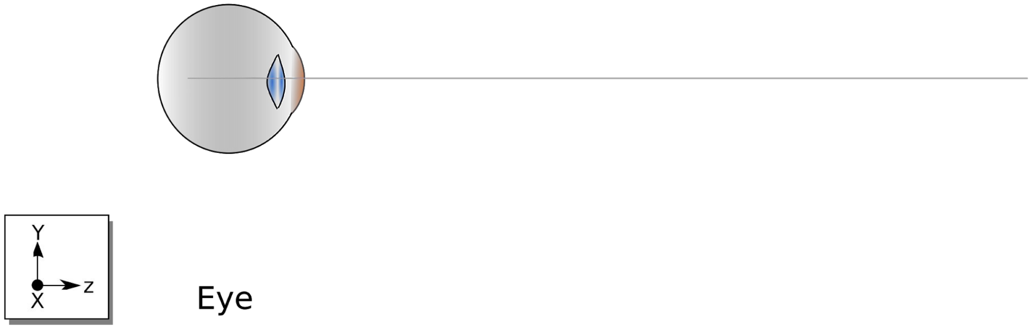
Real Max (2018)

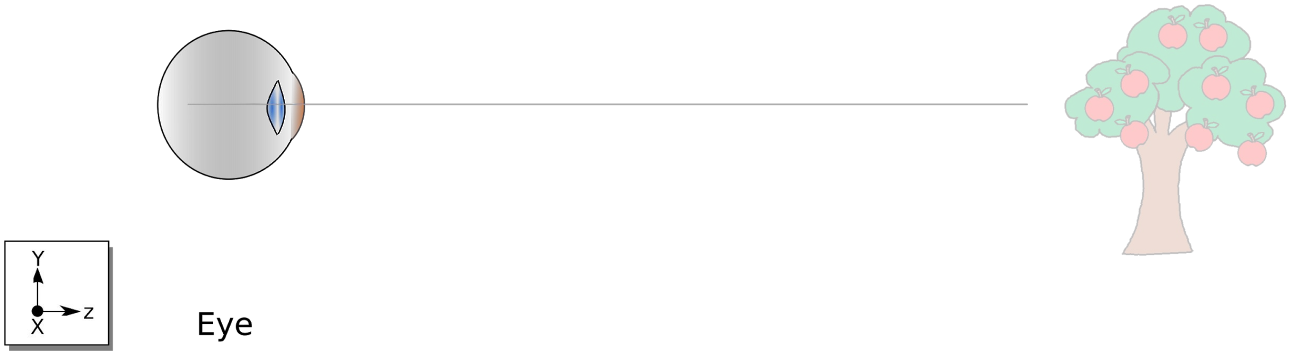
Kopin (2018)


Intel (2018)

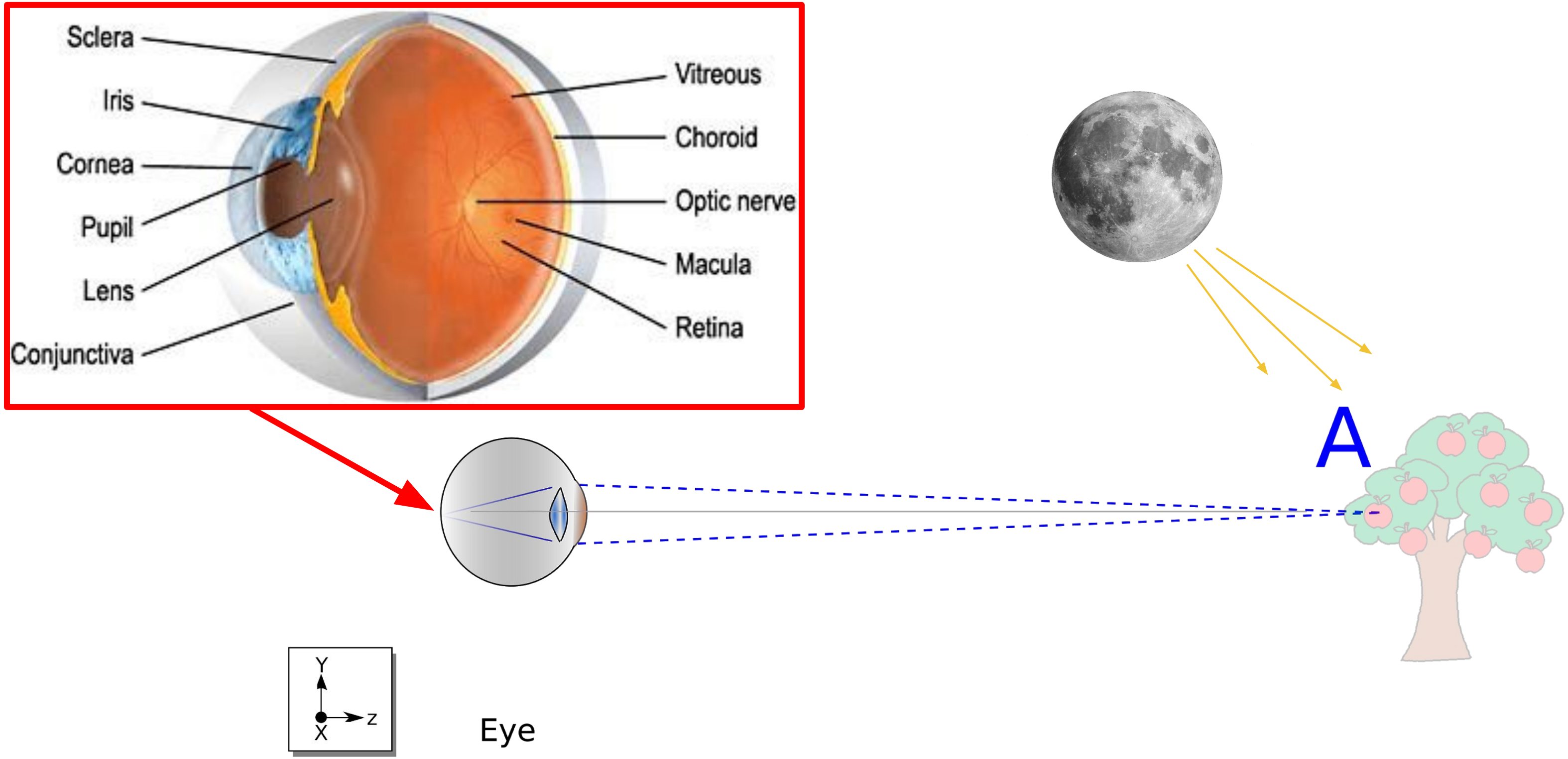
Magic Leap (2018)

How do they work?


Y
z
X
Eye

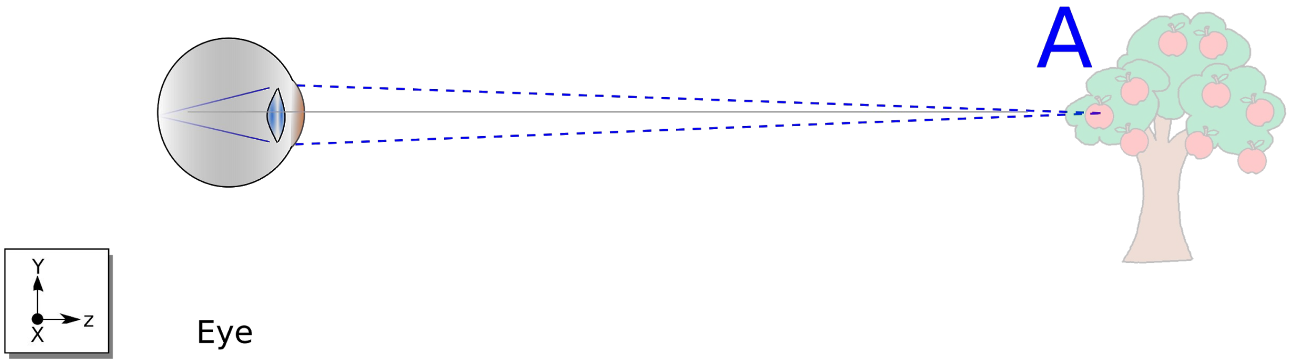


Y
Z
X
Eye

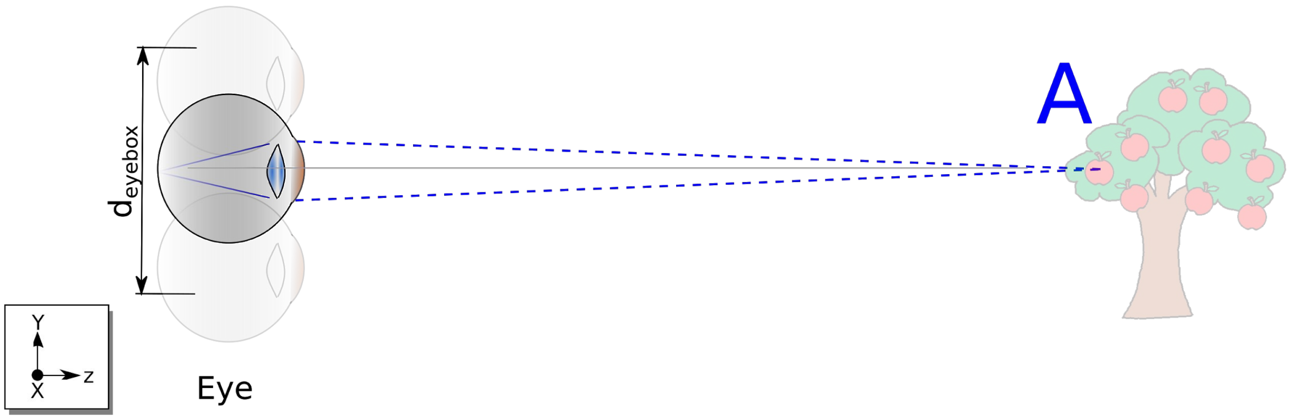


Real life is high dynamic range!

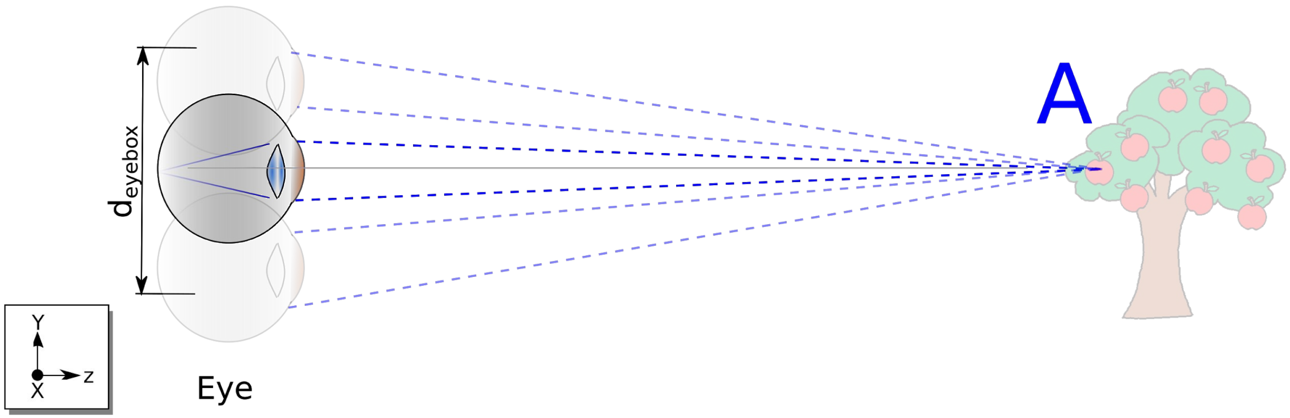
Reinhard, Erik, et al. *High dynamic range imaging: acquisition, display, and image-based lighting*. Morgan Kaufmann, 2010.


A
Y
Z
X
Eye

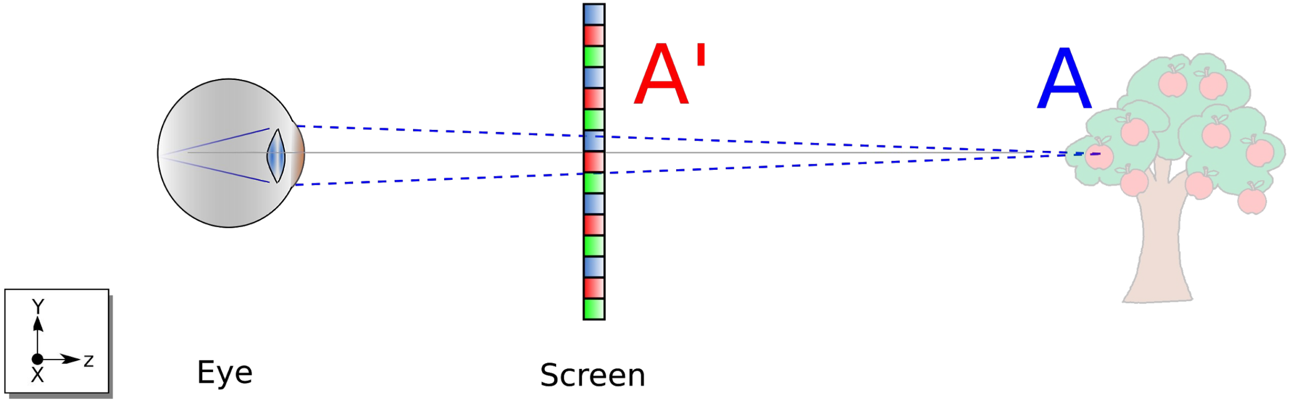
Real life has infinite eyebox/viewing zone!

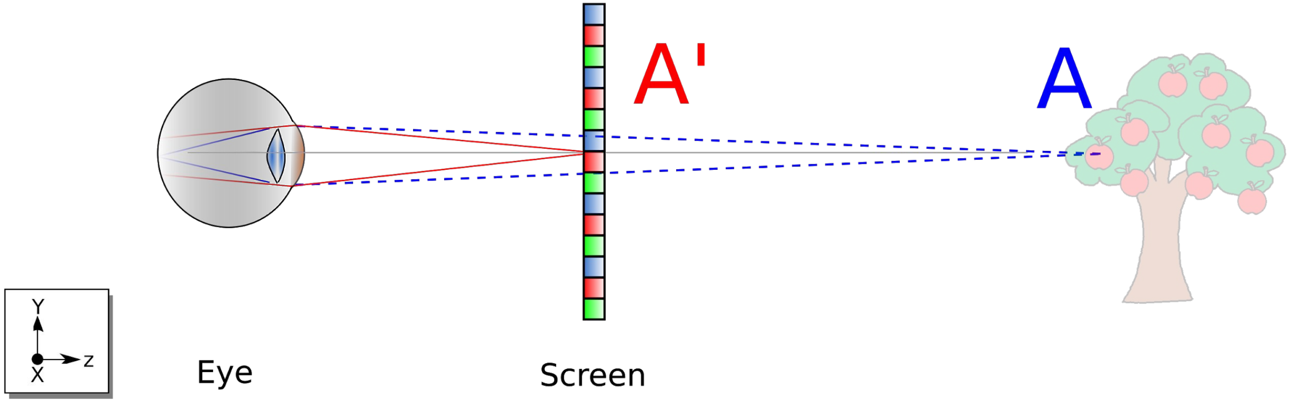


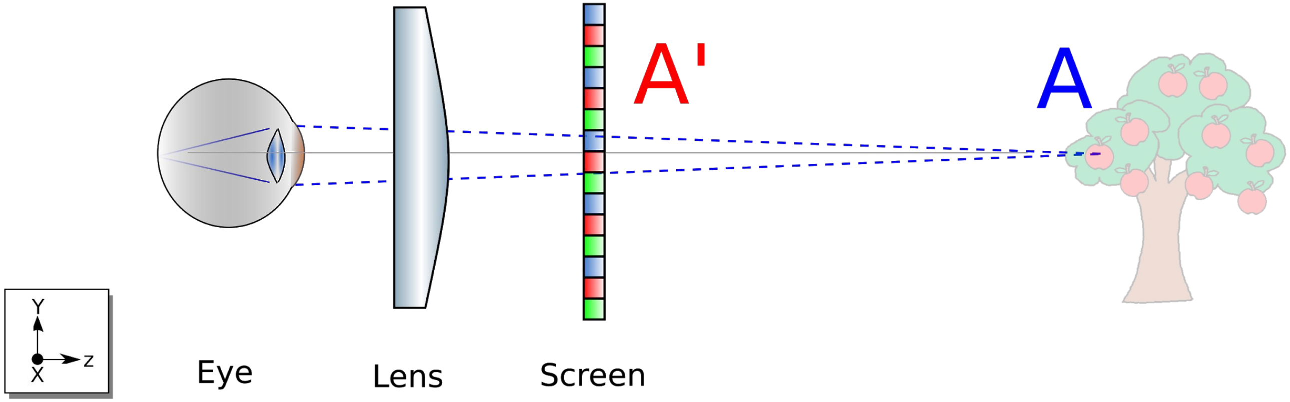
Real life is 4D Light Fields

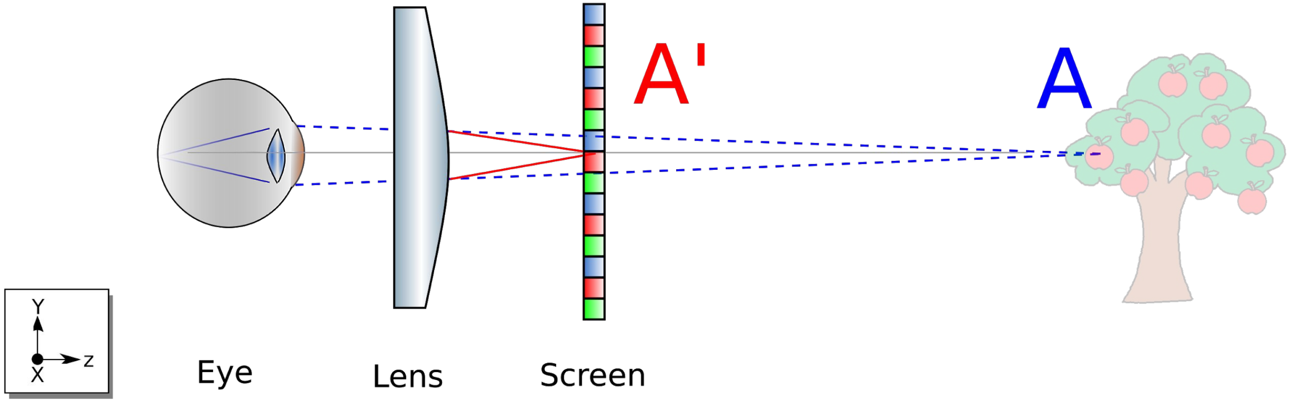


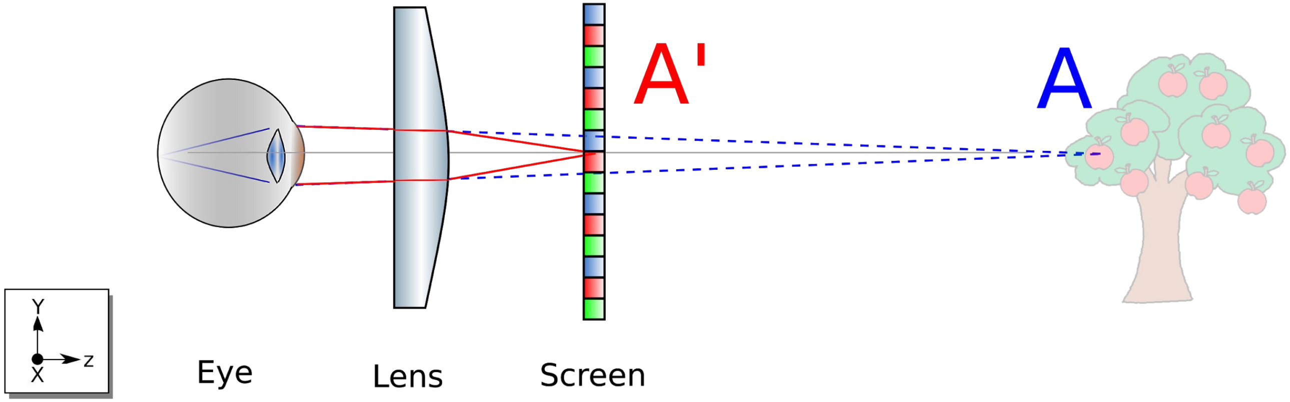
Levoy, Marc, and Pat Hanrahan. "Light field rendering." *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques*. ACM, 1996.

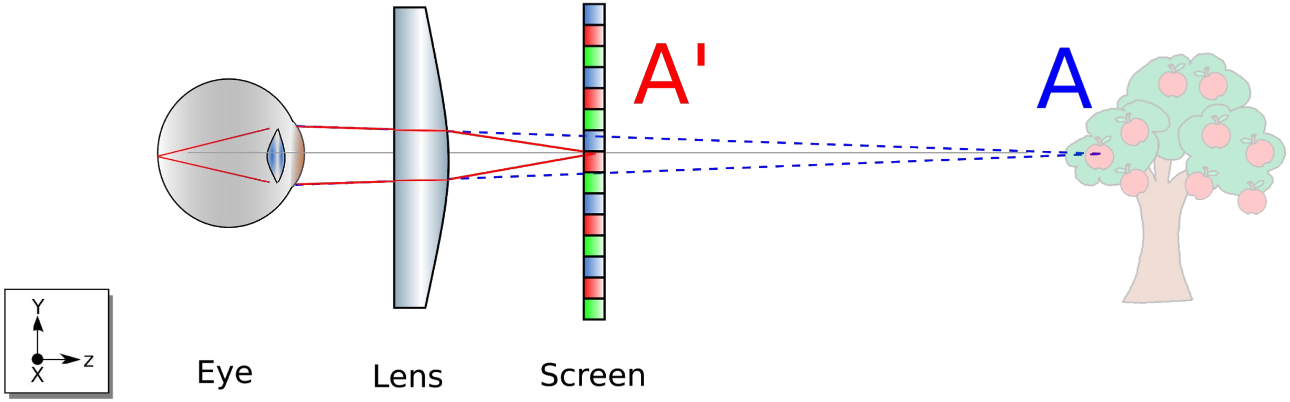

A'
A
Y
Z
X
Eye
Screen


A'
A
Y
Z
X
Eye
Screen

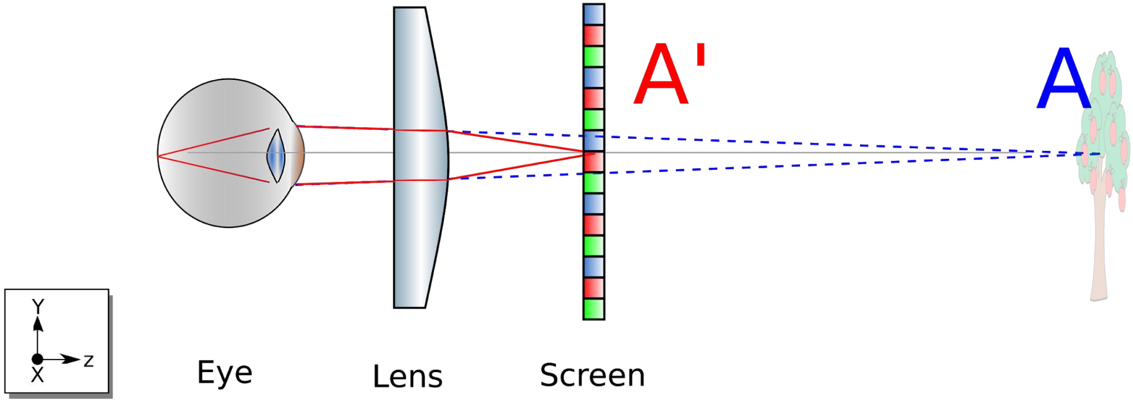

A'
A
Y
Z
X
Eye
Lens
Screen

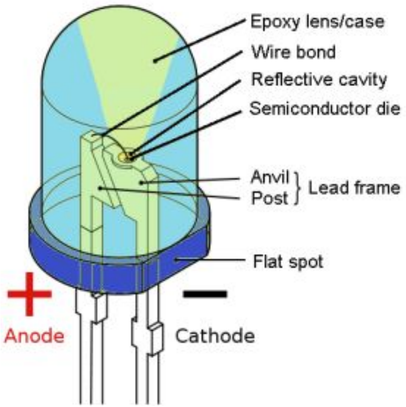

A'
A
Y
Z
X
Eye
Lens
Screen

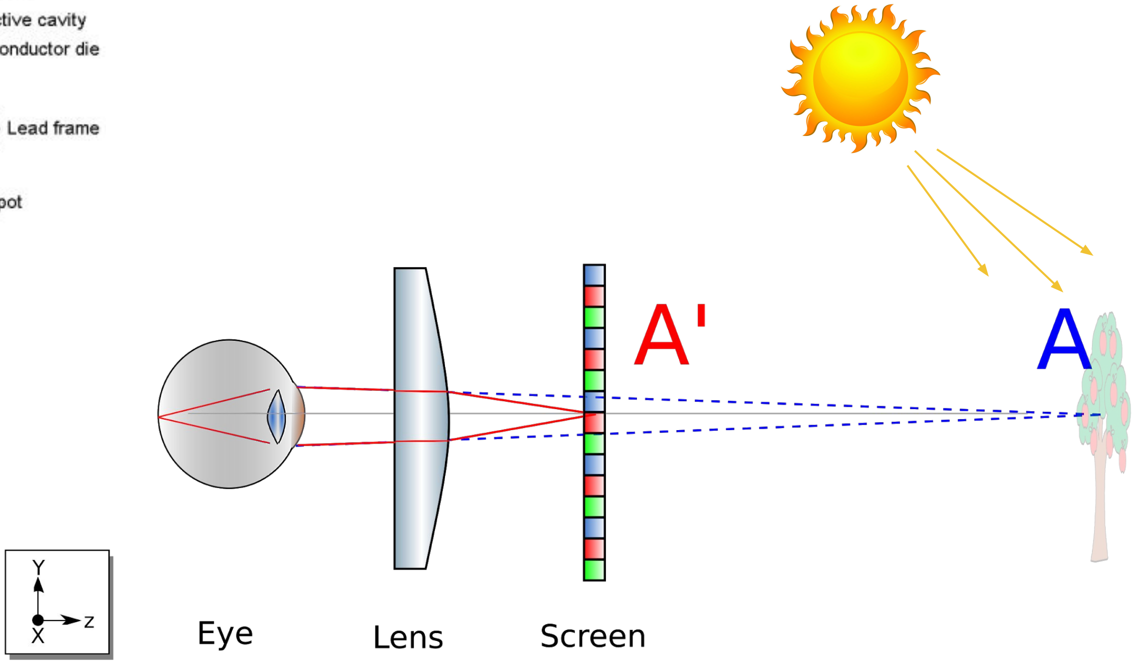

A'
A
Y
Z
X
Eye
Lens
Screen


A'
A
Y
Z
X
Eye
Lens
Screen

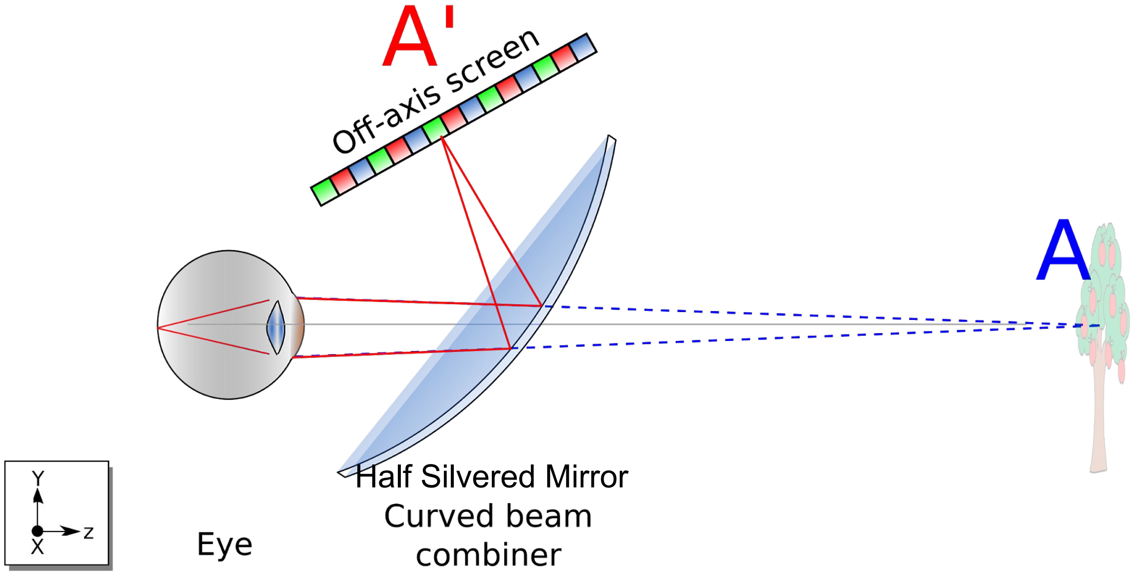
Current virtual reality near eye displays does not support different optical depth levels!





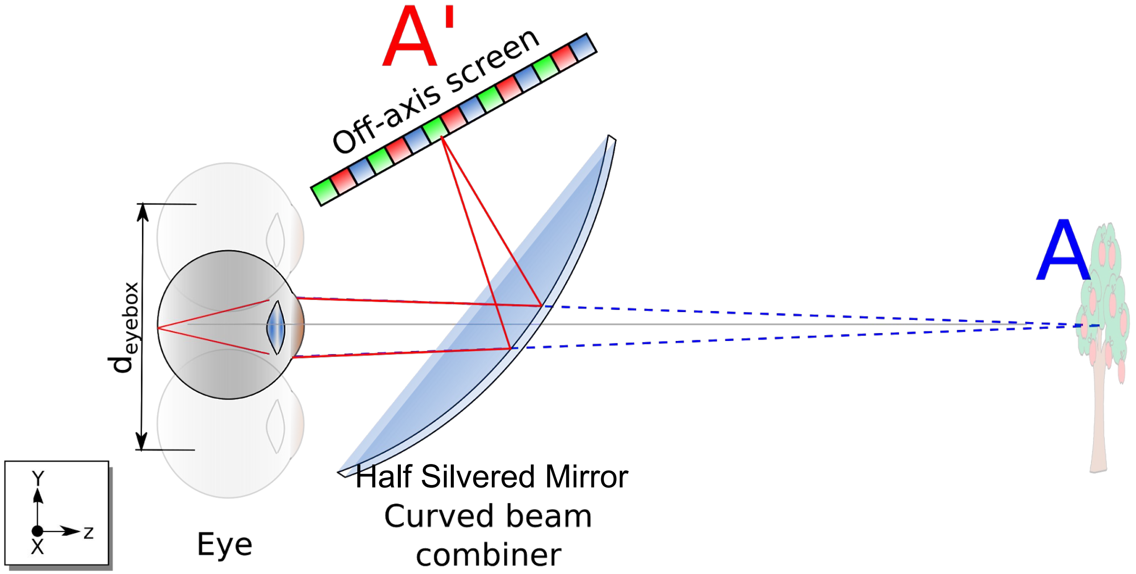


Current virtual reality near eye displays can not generate all the colors and can not support all brightness levels.



Pinhas Gilboa. 1991. Designing the right visor. In Medical Imaging. International Society for Optics and Photonics.

Current generation of augmented reality near eye displays can not generate wide eyebox as in the case of virtual reality near eye displays.



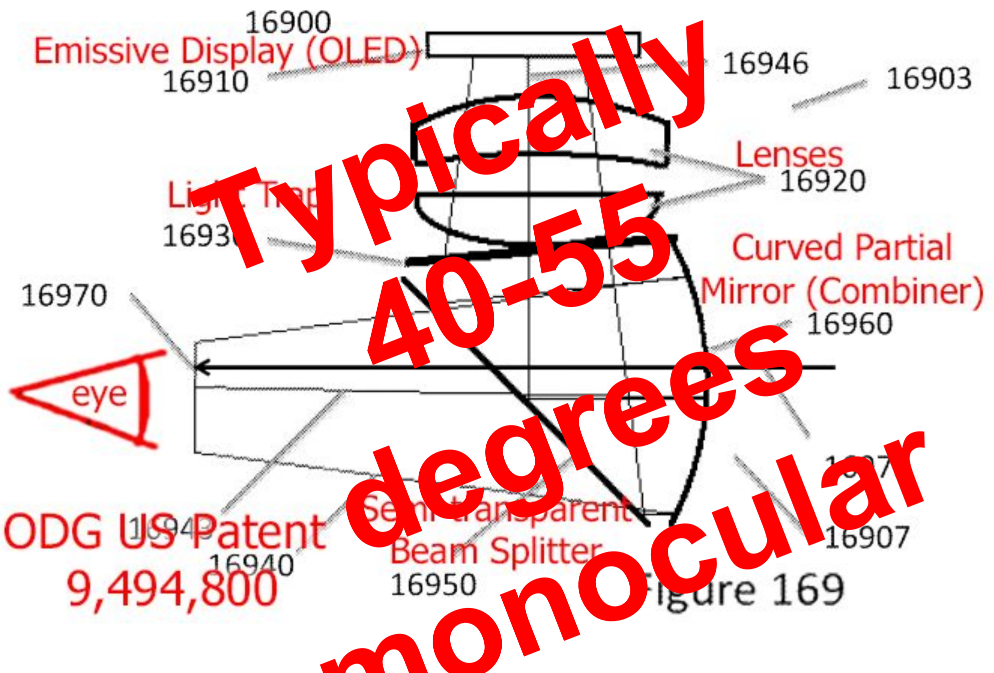
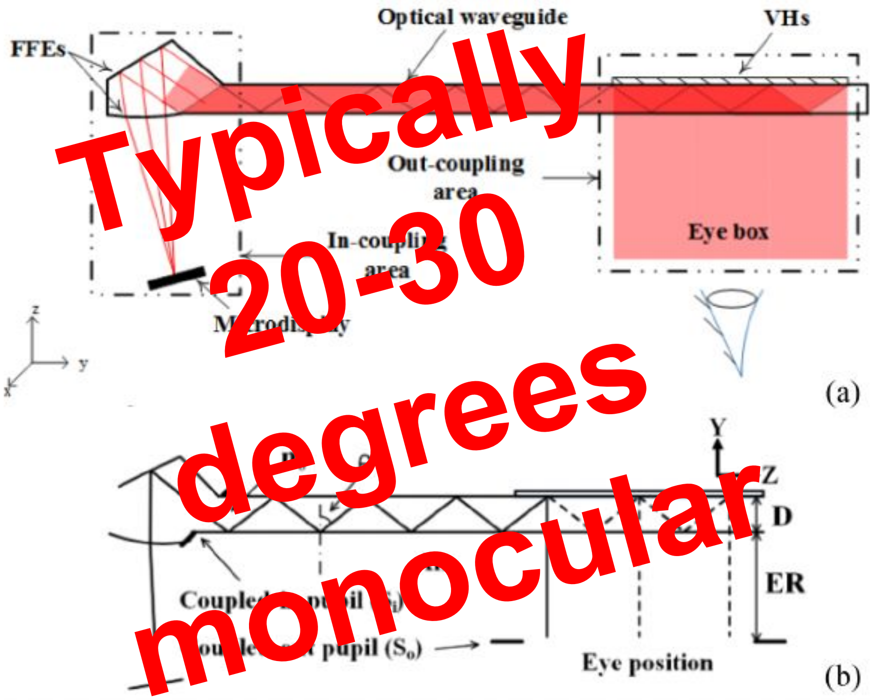


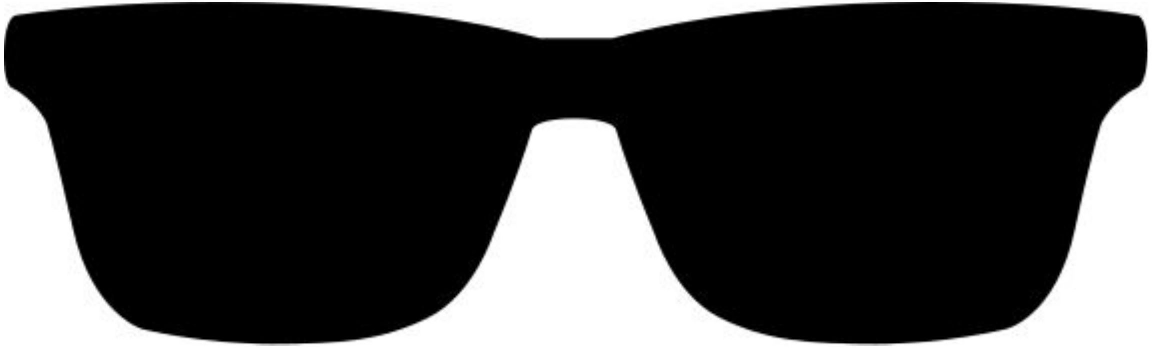
Image from
http://www.kguttag.com/2017/03/03/near-eye-bird-bath-optics-pros-and-cons-and-immys-different-approach/



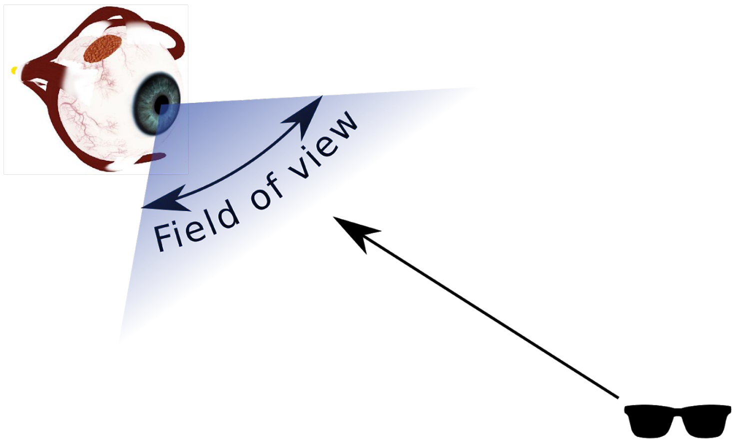
Han, Jian, et al. Optics express 23.3 (2015).

Current augmented reality near eye displays can not generate wide field of view.

Challenges?

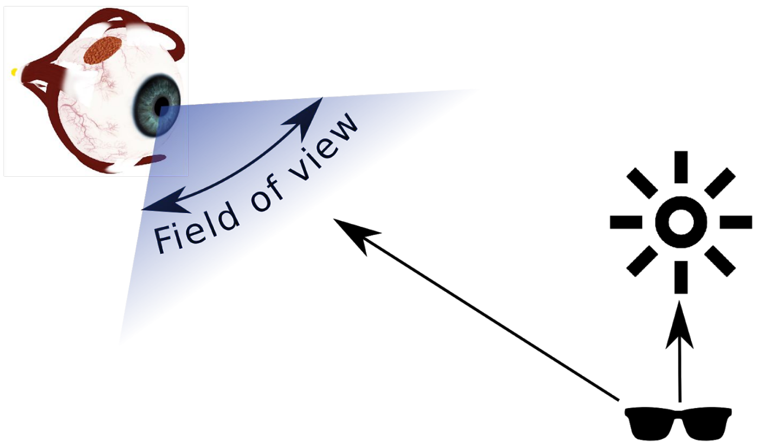


[Kramida, Gregory. IEEE transactions on visualization and computer graphics (2016),
Hua, Hong. Proceedings of the IEEE (2017)]

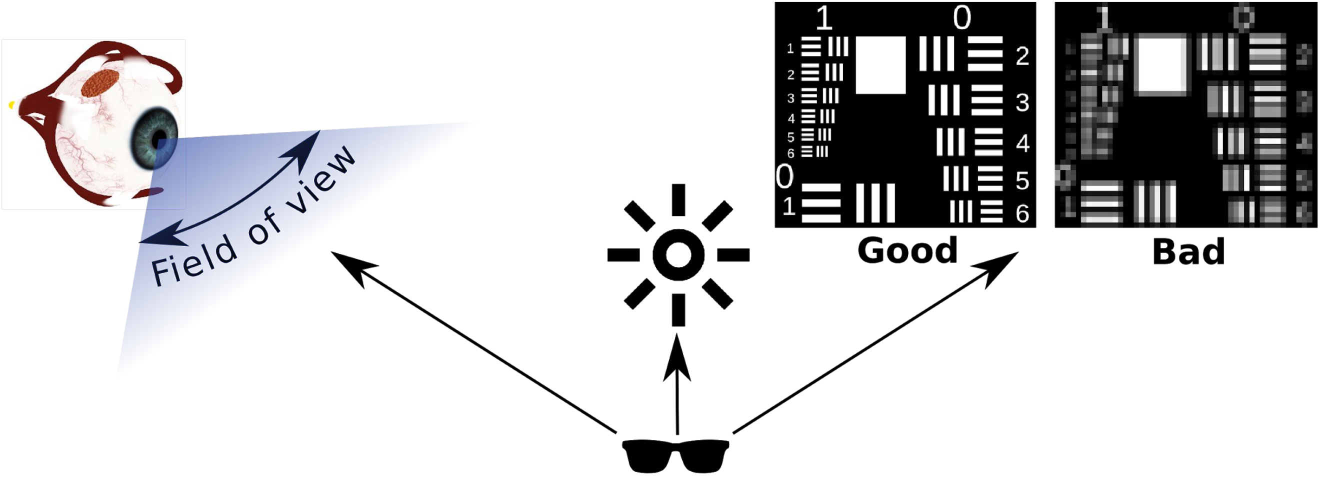


190 degrees of binocular field of view

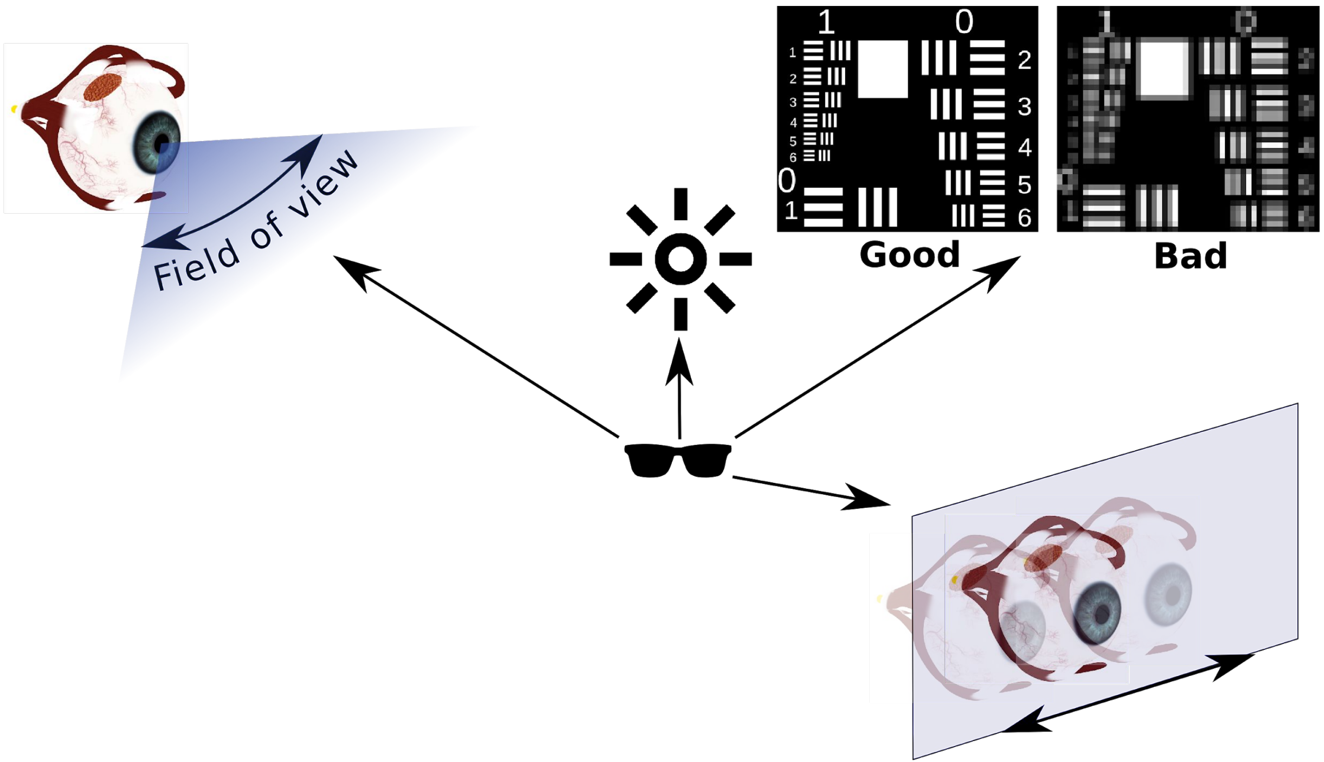
Paul Webb. 1964. Bioastronautics data book. (1964).



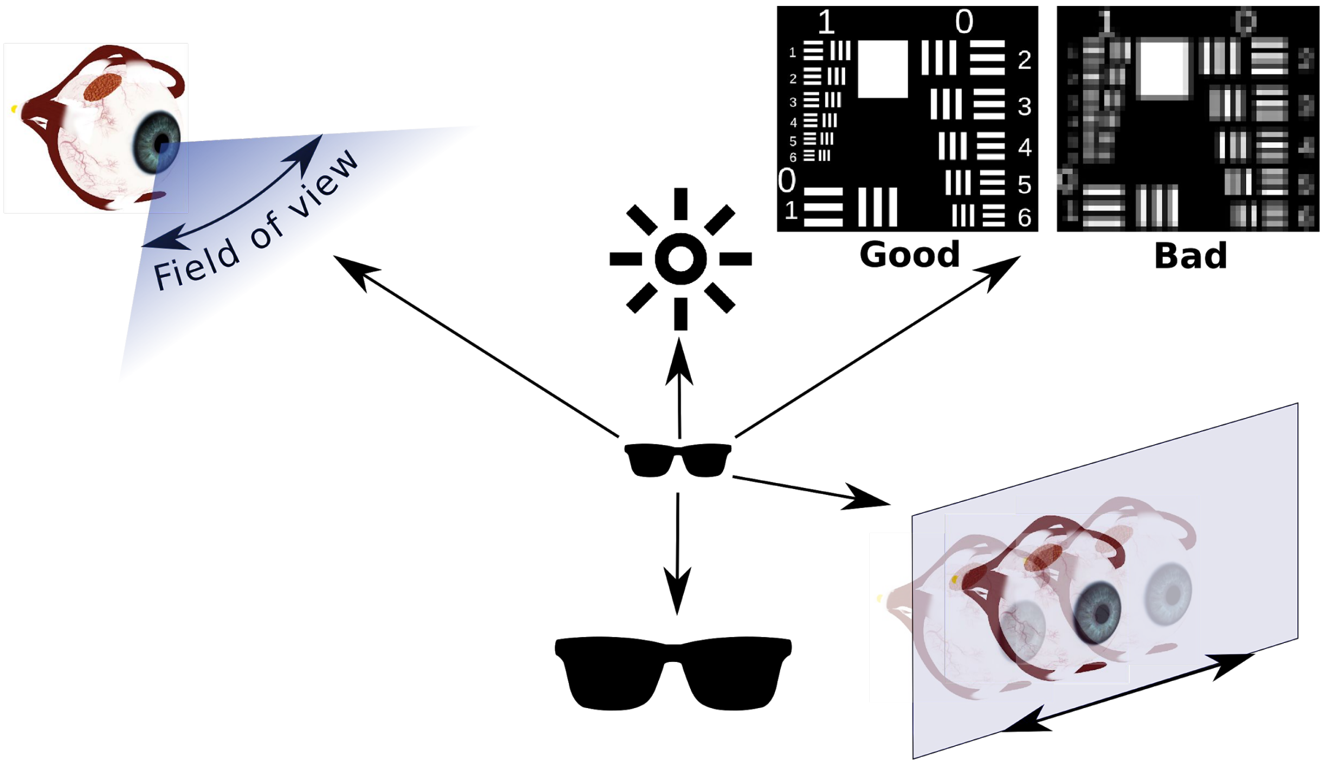
The human visual system can adapt from ~10^-6 cd/m^2 to ~10^6 cd/m^2. It has an unique color perception.



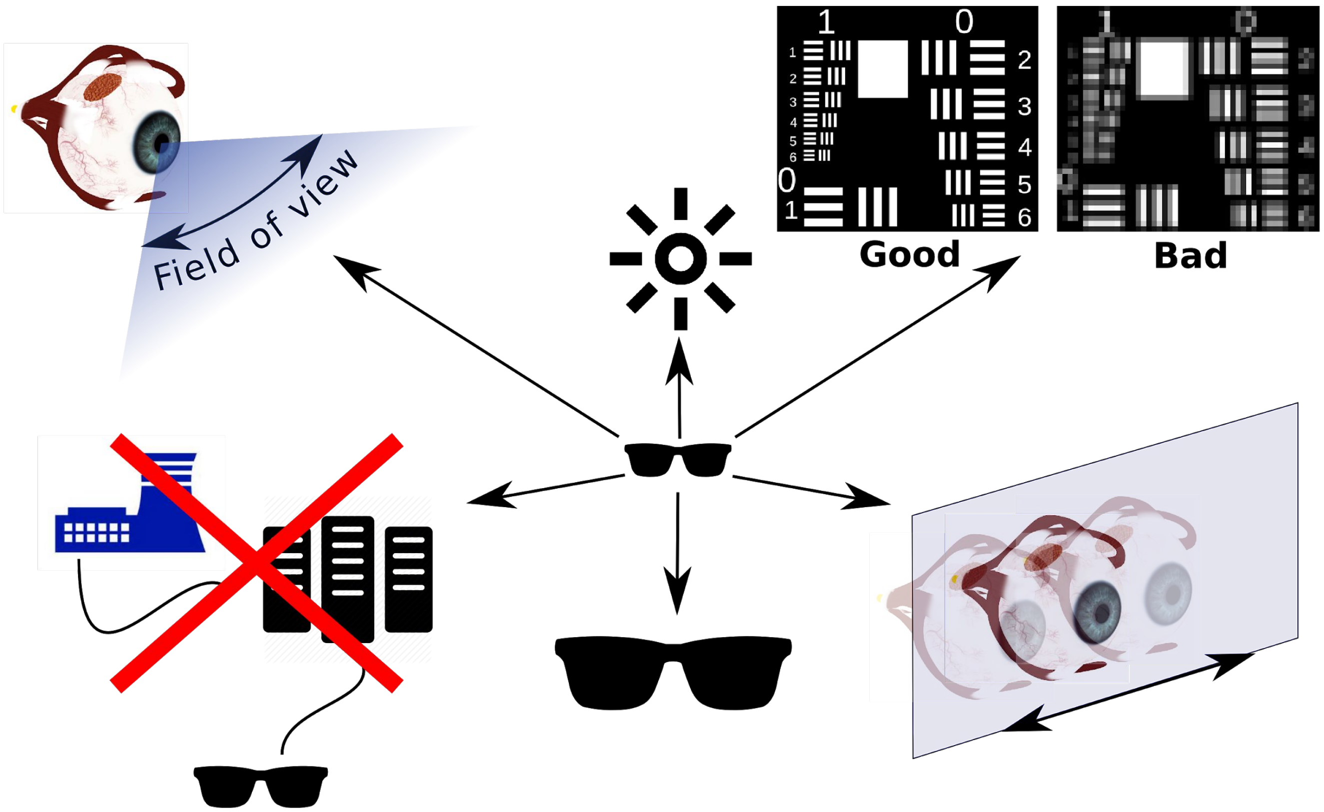
The human visual system has 20/20 visual acuity, 1 arcmin of resolution.



A large eyebox is needed in front of an eye,
typically 20 mm x 20 mm.

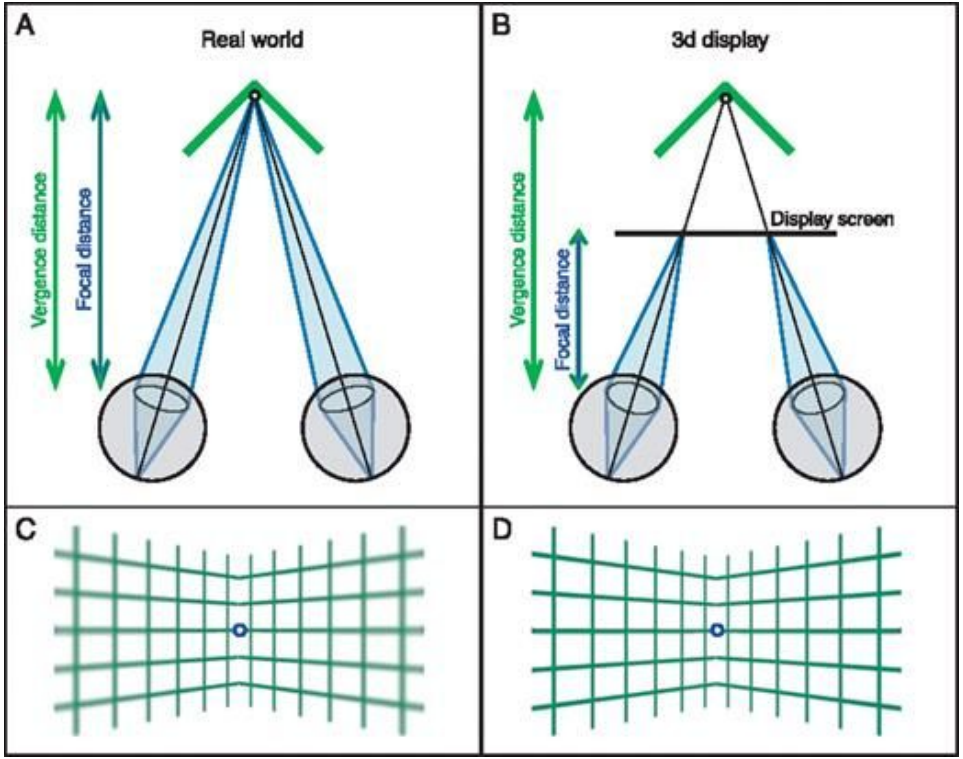


Slim form factor



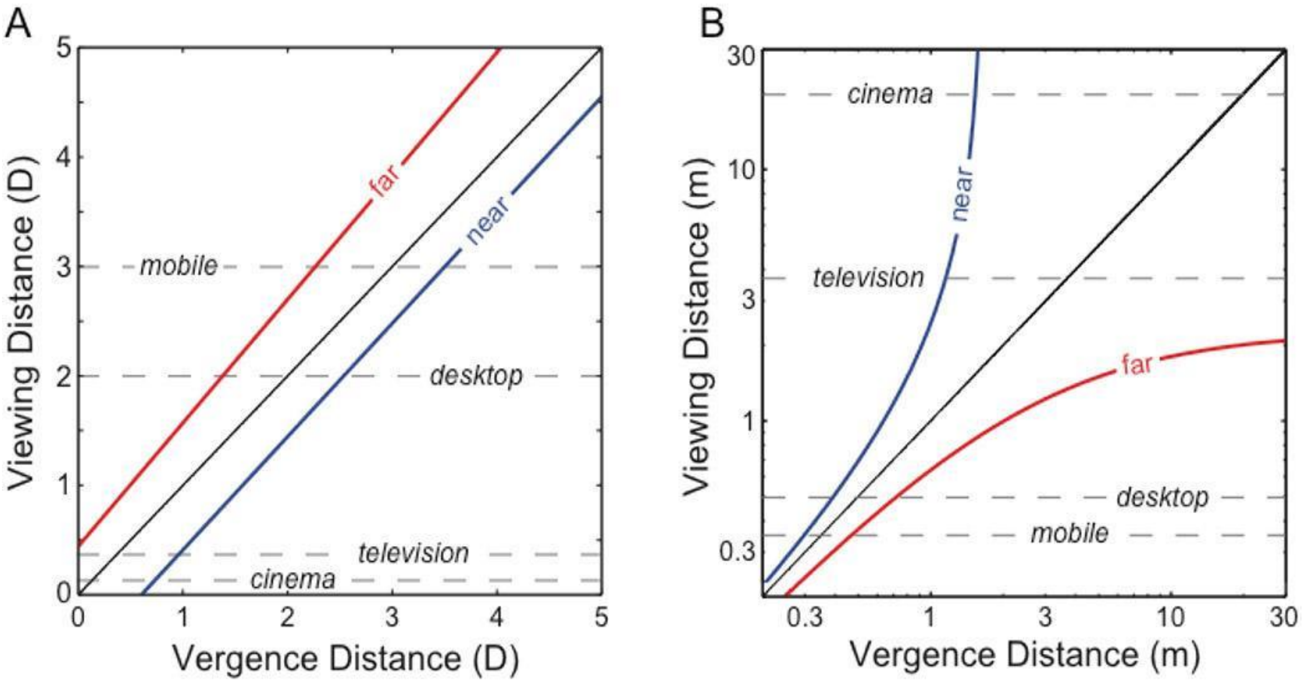
A typical smartphone has 5.45 Wh energy
with an 1.7Ghz Quad-Core ARM Cortex-A53
CPU.

Accommodation - Vergence Conflict



[Hoffman, David M., et al. Journal of vision 8.3 (2008): 33-33.]

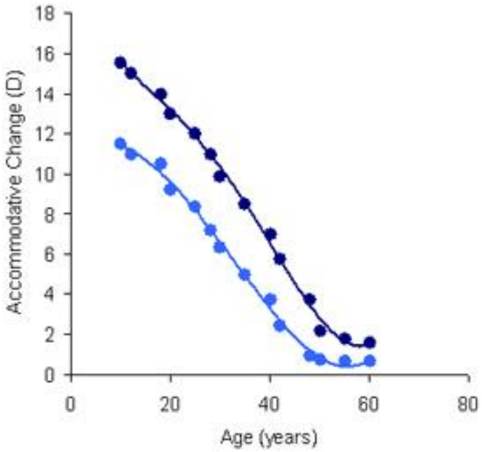
Zone of Comfort



[T. Shibata, et al *Journal of vision* (2011)]

Presbyopia

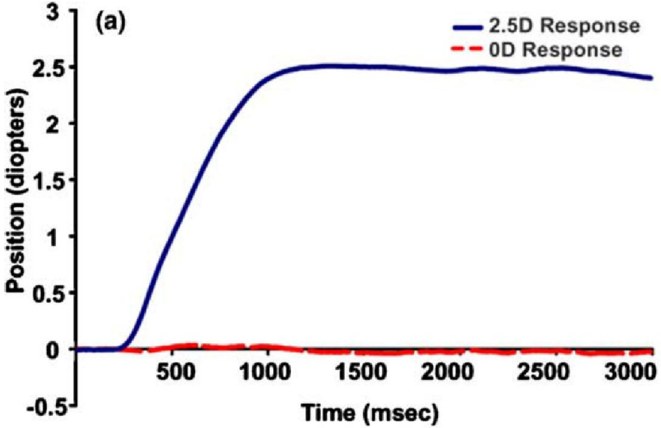
- As we age, our focal adaptation weakens
- For those advanced in age, having fixed focus in VR can be good if it is the right focus
- Not so for optical see-through AR: when the real world needs to be corrected



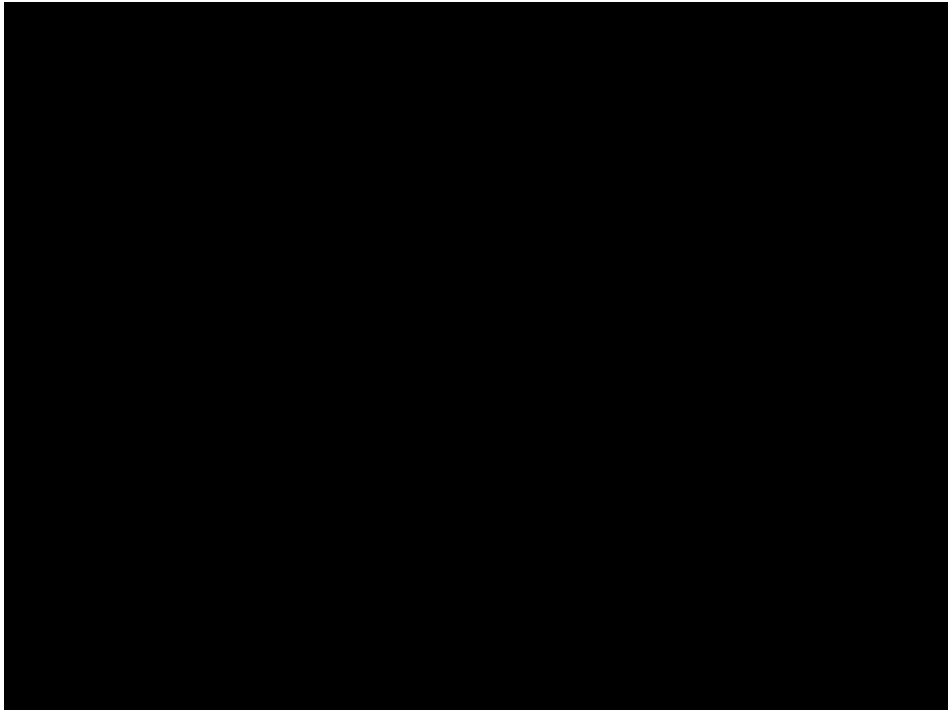
http://www.cvs.rochester.edu/yoonlab/research/pa.html
http://eyeglasses-asheville.com

Accommodation response

- Step change of fixated object depth
  - Smooth and steady accommodation increase
    - up to 1 second to achieve the full accommodation state
    - ~300 ms latency



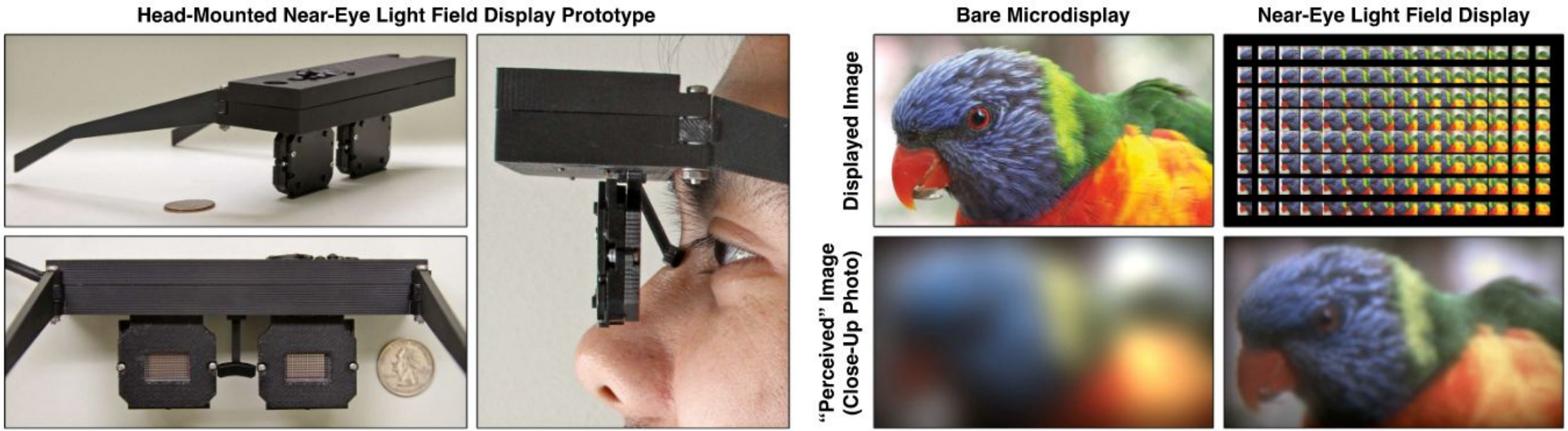
[Bharadwaj and Schor, Vision Research 2004]

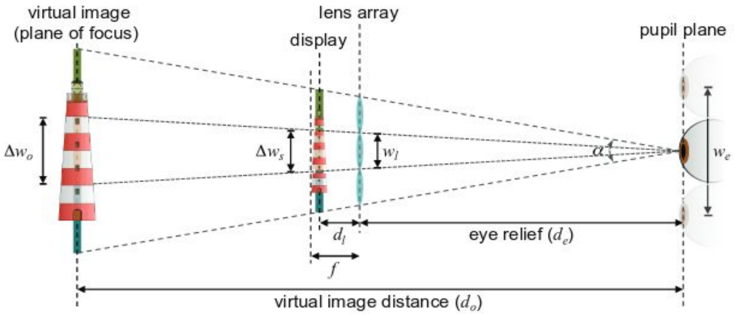


Video from Edmund Optics

Investment : >1-5 Million USD + Permanent technical personnel + Long processing times (6-8 weeks)

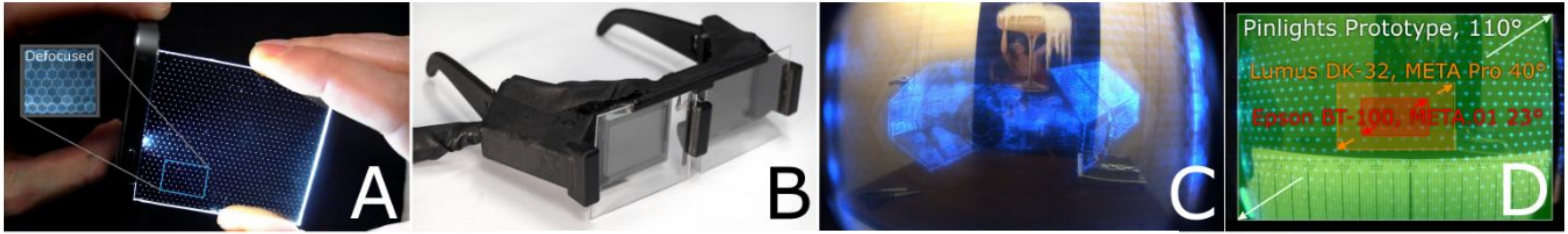
Nvidia’s near eye displays

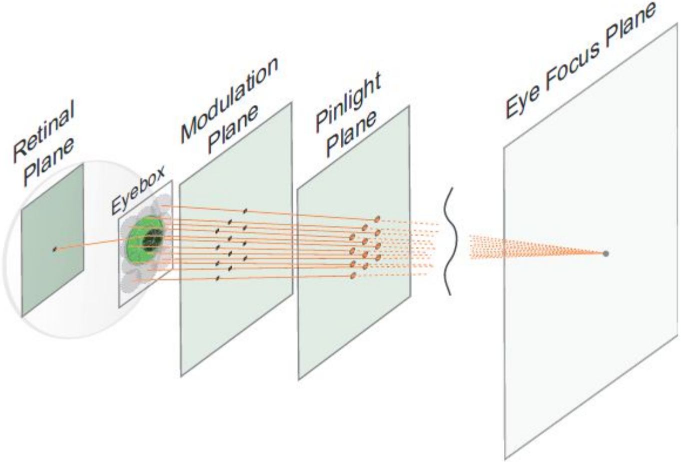




Microlens displays

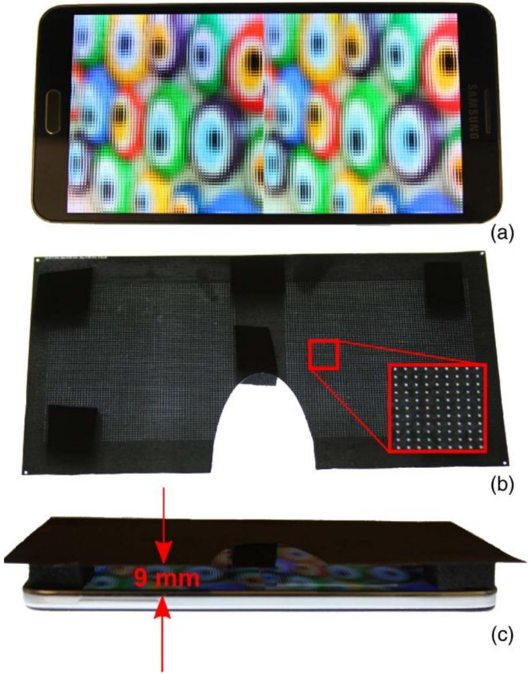
[Lanman and Luebke *ACM SIGGRAPH ASIA* 2013]

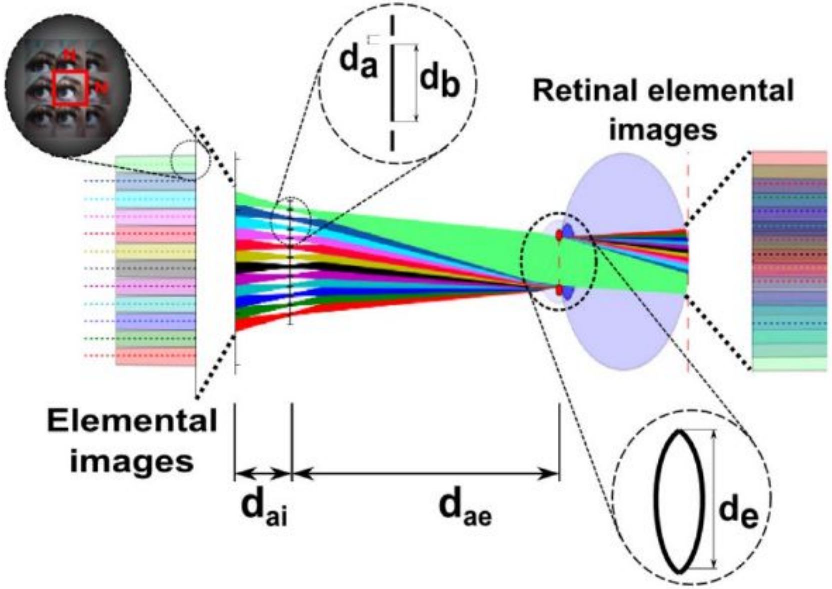




Pinlight displays

[Maimone et al. *ACM SIGGRAPH 2014]*

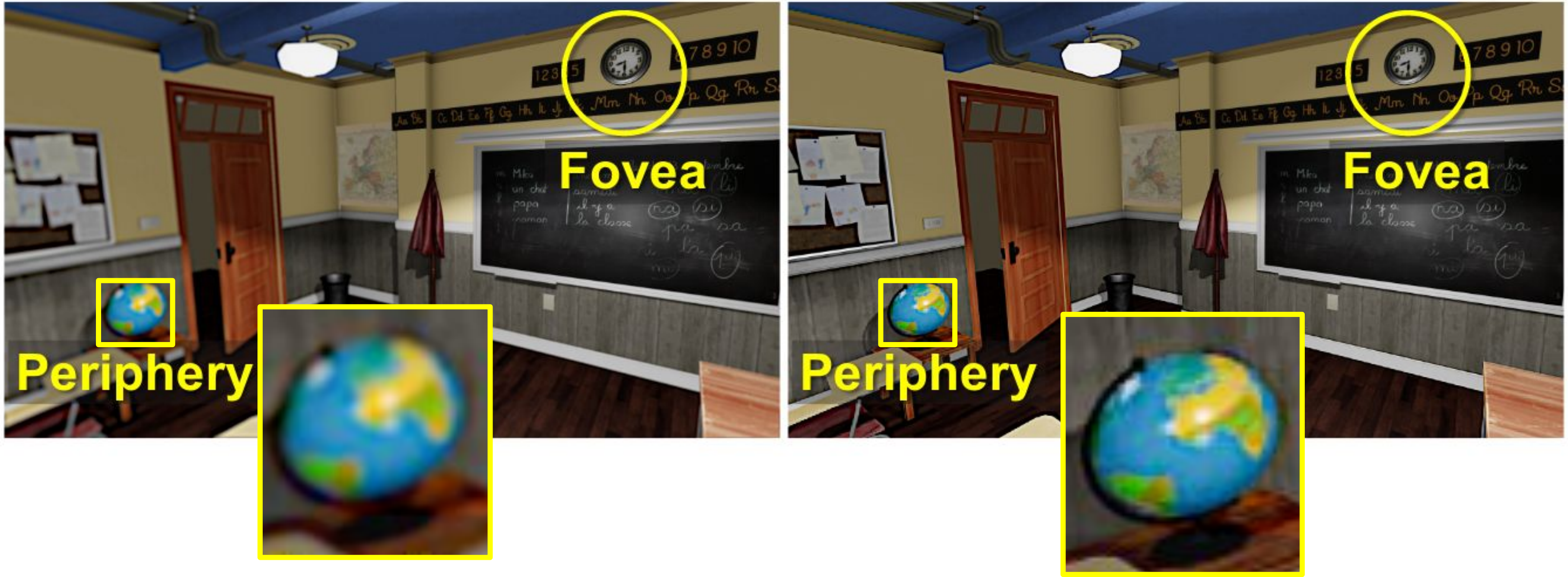




Pinhole displays

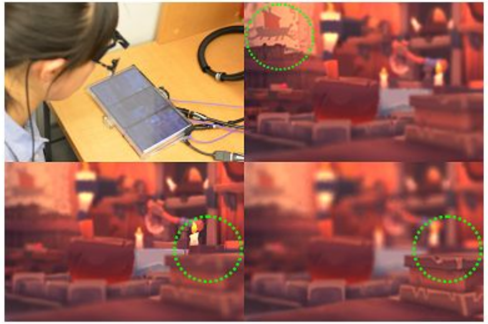
[Kaan Akşit et al. *Applied optics, 2015*]

NEED GAZE AWARE RENDERING



Patney et al."Perceptually-based foveated virtual reality." In *ACM SIGGRAPH 2016 Emerging Technologies*, p. 17. ACM, 2016.

Perceptually-Guided Foveation for Light Field Displays



A variety of applications such as virtual reality and immersive cinema require high image quality, low rendering latency, and consistent depth cues. 4D light field displays support focus accommodation, but are more costly to render than 2D images, resulting in higher latency. The human visual system can resolve higher spatial frequencies in the fovea than in the periphery. This property has been harnessed by recent 2D foveated rendering methods to reduce computation cost while maintaining perceptual quality. Inspired by this, we present foveated 4D light fields by investigating their effects on 3D depth perception. Based on our psychophysical experiments and theoretical analysis on visual and display bandwidths, we formulate a content-adaptive importance model in the 4D ray space. We verify our method by building a prototype light field display that can render only 16%-30% rays without compromising perceptual quality.

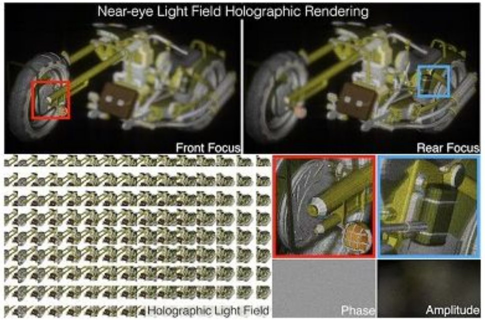
Authors: Qi Sun (Stony Brook University & NVIDIA)
Fu-Chung Huang
Joohwan Kim
Li-Yi Wei (University of Hong Kong)
David Luebke
Arie Kaufman (Stony Brook University)

Publication Date: Monday, November 27, 2017

Published in: ACM SIGGRAPH ASIA 2017

[Qi et al., *ACM SIGGRAPH 2017]*

Near-eye Light Field Holographic Rendering with Spherical Waves for Wide Field of View Interactive 3D Computer Graphics



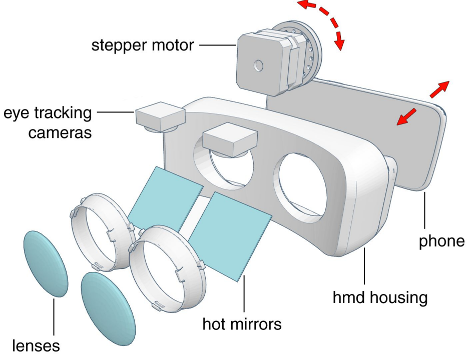
Holograms have high resolution and great depth of field allowing the eye to view a scene much like seeing through a virtual window. Unfortunately, computer generated holography (CGH) does not deliver the same promise due to hardware limitations under plane wave illumination and large computational cost. Light field displays have been popular due to their capability to provide continuous focus cue. However, light field displays suffer from the trade offs between spatial and angular resolution, and do not model diffraction. We present a light field based CGH rendering pipeline allowing for reproduction of high-definition 3D scenes with continuous depth and support of intra-pupil view dependent occlusion. Our rendering accurately accounts for diffraction and supports various types of reference illumination for holograms. We prevent under- and over-sampling and geometric clipping suffered in previous work. We also implement point-based methods with Fresnel integration that are orders of magnitude faster than the state of art, achieving interactive volumetric 3D graphics. To verify our computational results, we build a see-through near-eye color display prototype with CGH that enables co-modulation of both amplitude and phase. We show that our rendering accurately models the spherical illumination introduced by the eye piece and produces the desired 3D imaginary at designated depth. We also derive aliasing, theoretical resolution limits, depth of field, and other design trade-off space for near-eye CGH.

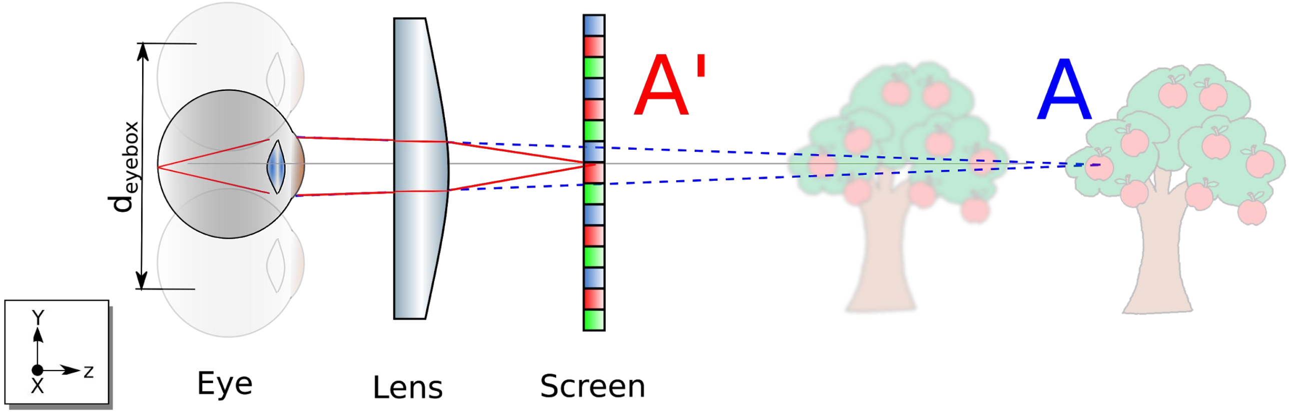
Authors: Liang Shi (NVIDIA & MIT CSAIL)
Fu-Chung Huang
Ward Lopes
Wojciech Matusik (MIT CSAIL)
David Luebke

[Liang et al. *Siggraph Asia, 2017*]

Varifocal display proposal I

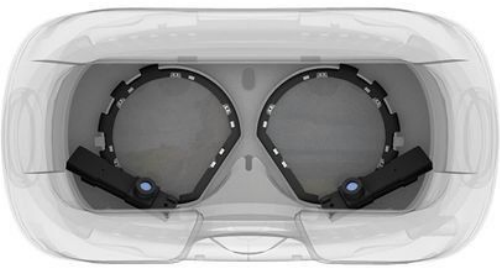
Kaan Akşit, Ward Lopes, Jonghyun Kim, Peter Shirley, and David Luebke. 2017. Near-eye varifocal augmented reality display using see-through screens. *ACM Trans. Graph.* 36, 6, Article 189 (November 2017)



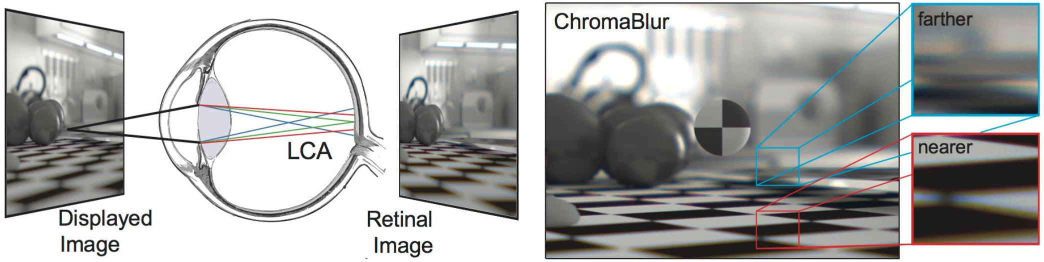


Our understanding of varifocal is aligned with

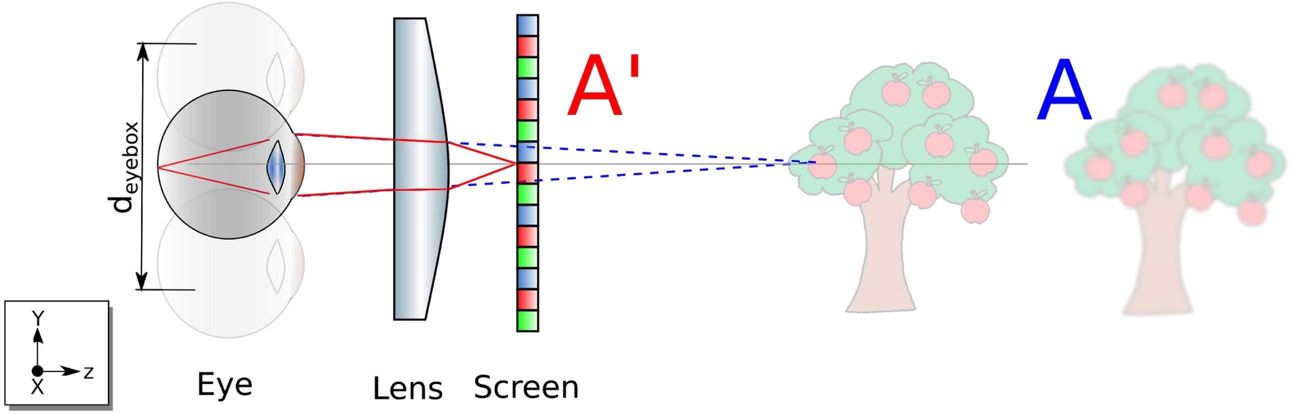
Padmanaban, Nitish, et al. "Optimizing virtual reality for all users through gaze-contingent and adaptive focus displays." Proceedings of the National Academy of Sciences (2017): 201617251.



Pupillabs eye tracker for HTC Vive



Cholewiak, Steven A., et al. "ChromaBlur: Rendering chromatic eye aberration improves accommodation and realism." Siggraph Asia (2017).



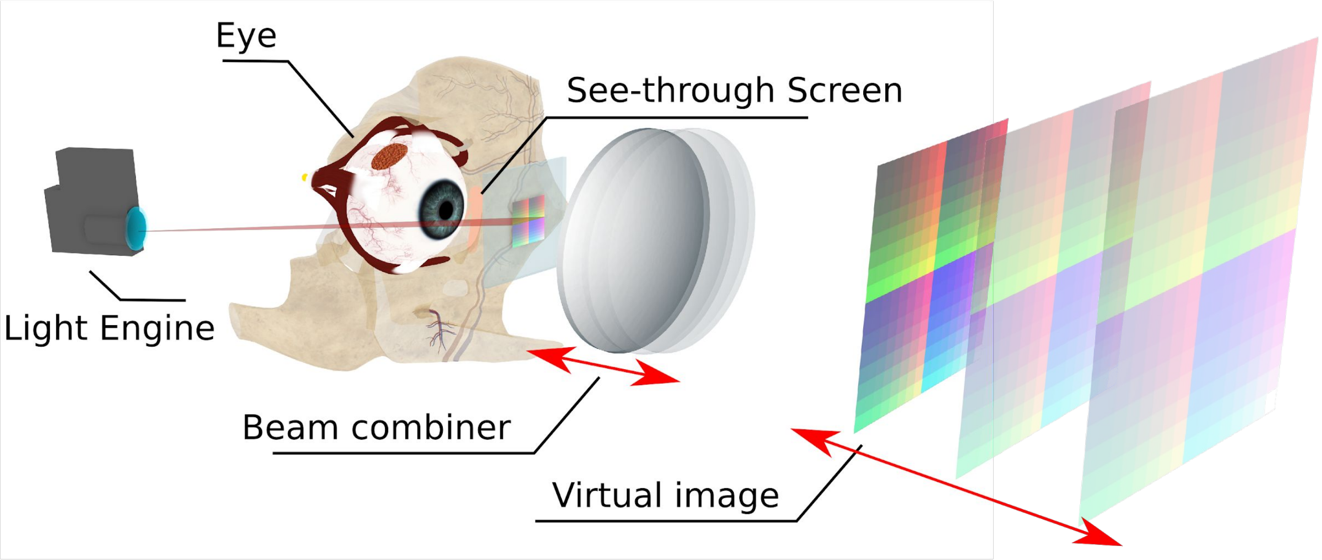
Moving depth plane in synchronism with an eye tracker, and applying a computational blur for mimicking optical blur.

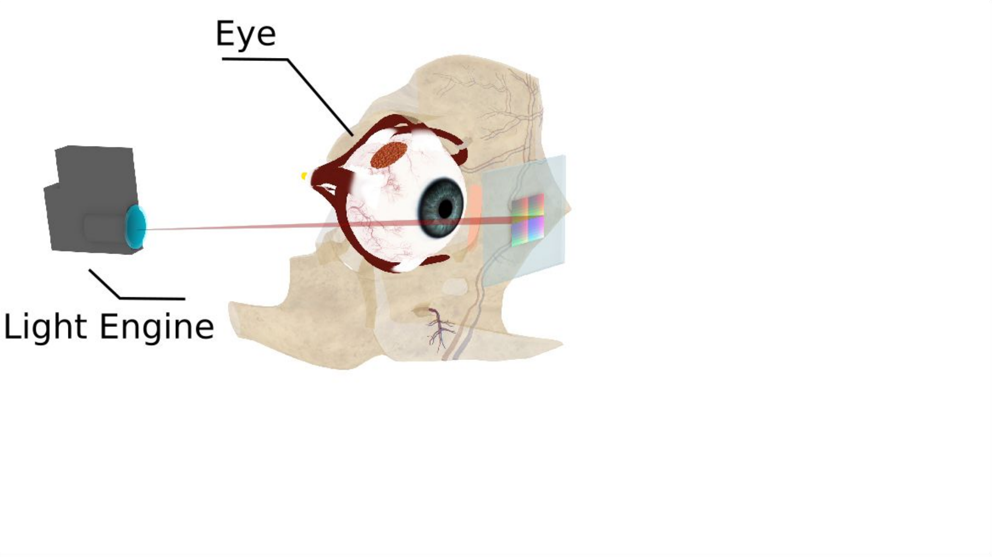
"Studies show evidence that supporting accommodative cues through a varifocal mechanism improves visual comfort and user performance while being simpler than other methods, but most current approaches sacrifice FoV and bulk."

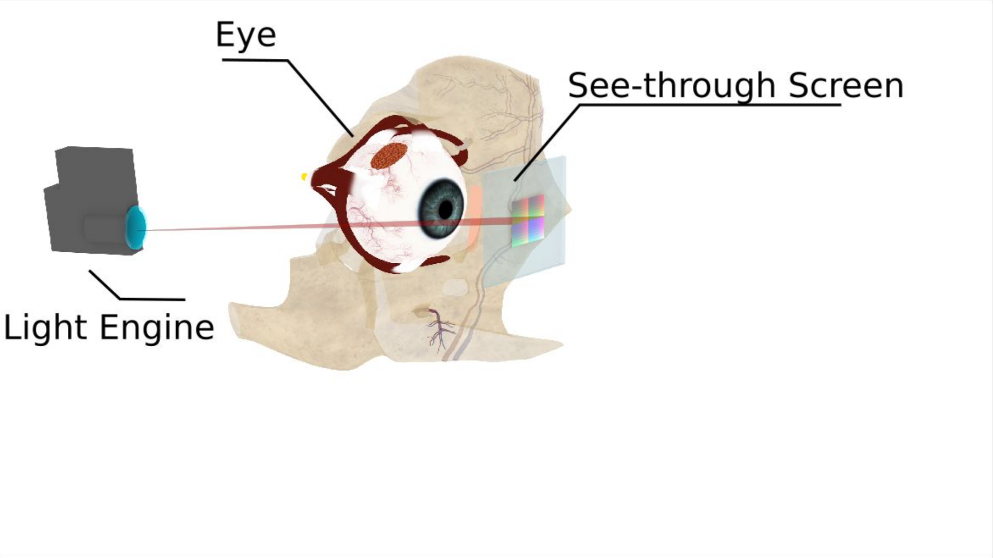
[Johnson et al. Optics Express 2016, Konrad et al. Human Factors in Computing 2016]

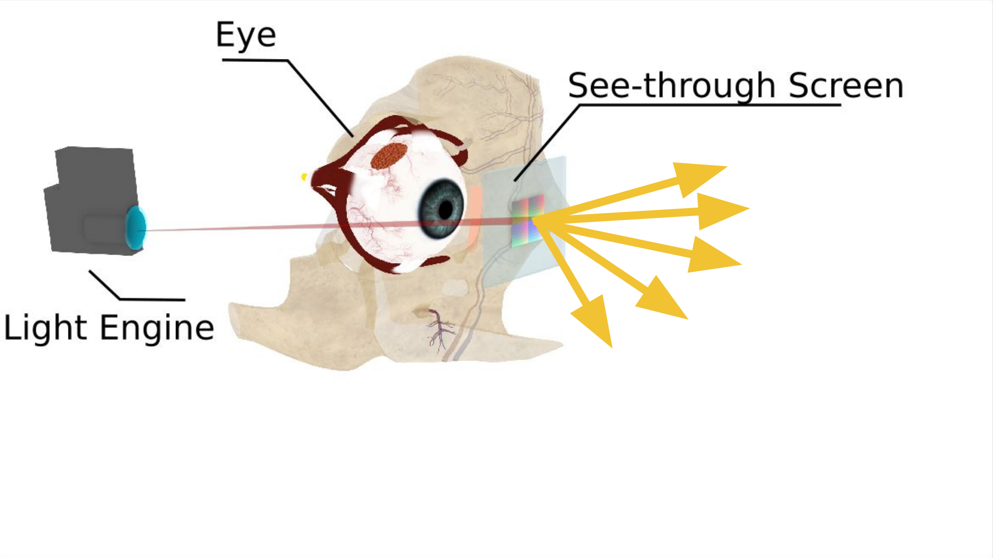
"The duration of actual lens accommodation of 500−800 ms has been reported, which means that the complete accommodation cycle, including the latency, typically requires around 1 second."

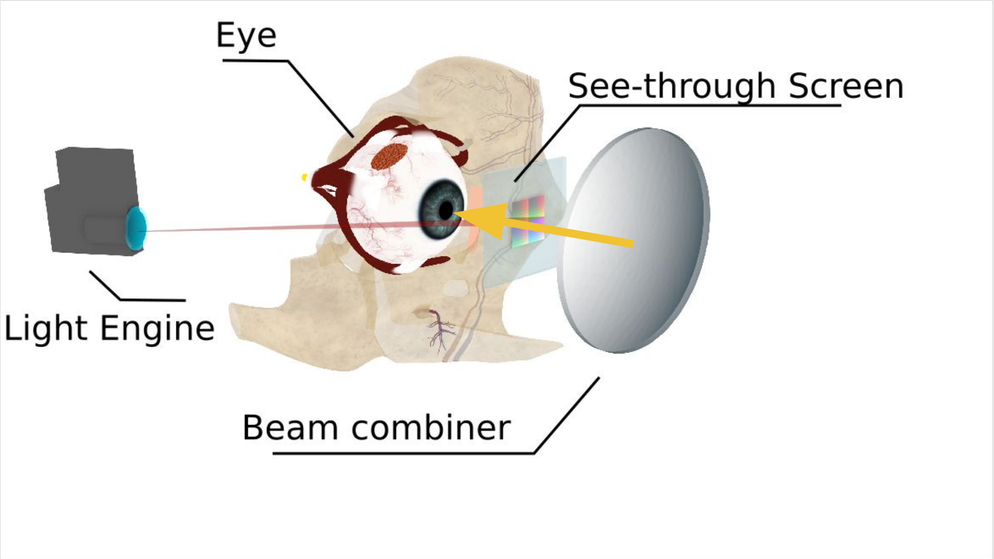
[S. R. Bharadwaj and C. M. Schor. Vision Research, (2005), F. Campbell and G. Westheimer. J. Physiol., (1960), G. Heron, W. Charman, and C. Schor. Vision Research, (2001), P. S., D. Shirachi, and S. L. American Journal of Optometry & Archives of American Academy of Optometry, (1972)]

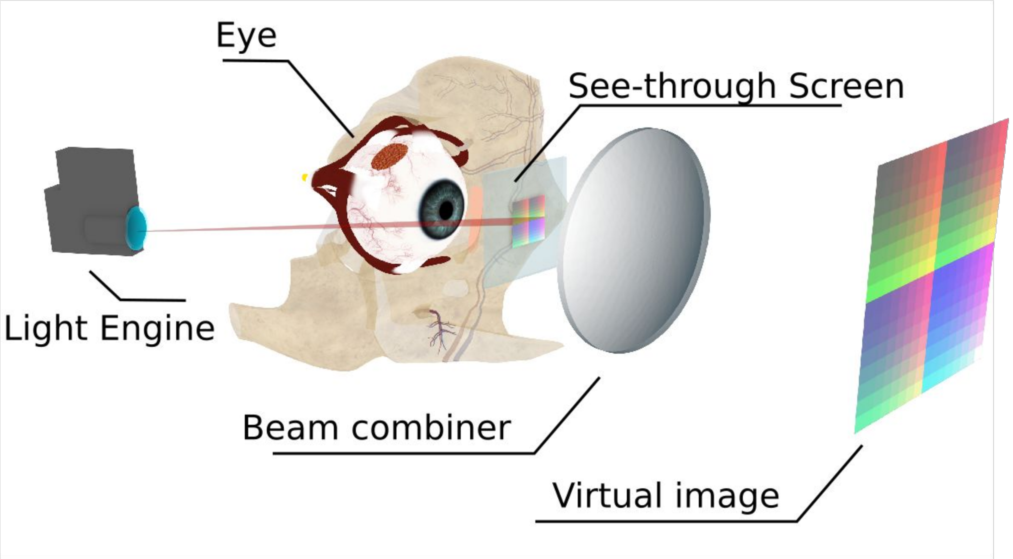

Eye
See-through Screen
Light Engine
Beam combiner
Virtual image

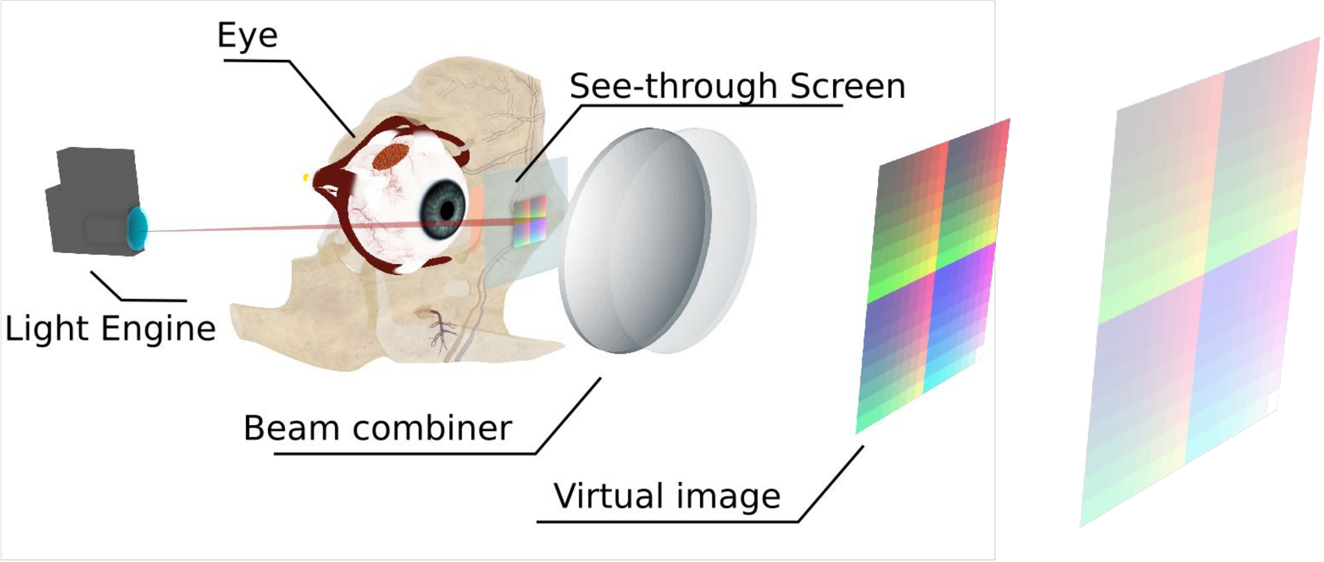

Eye
Light Engine

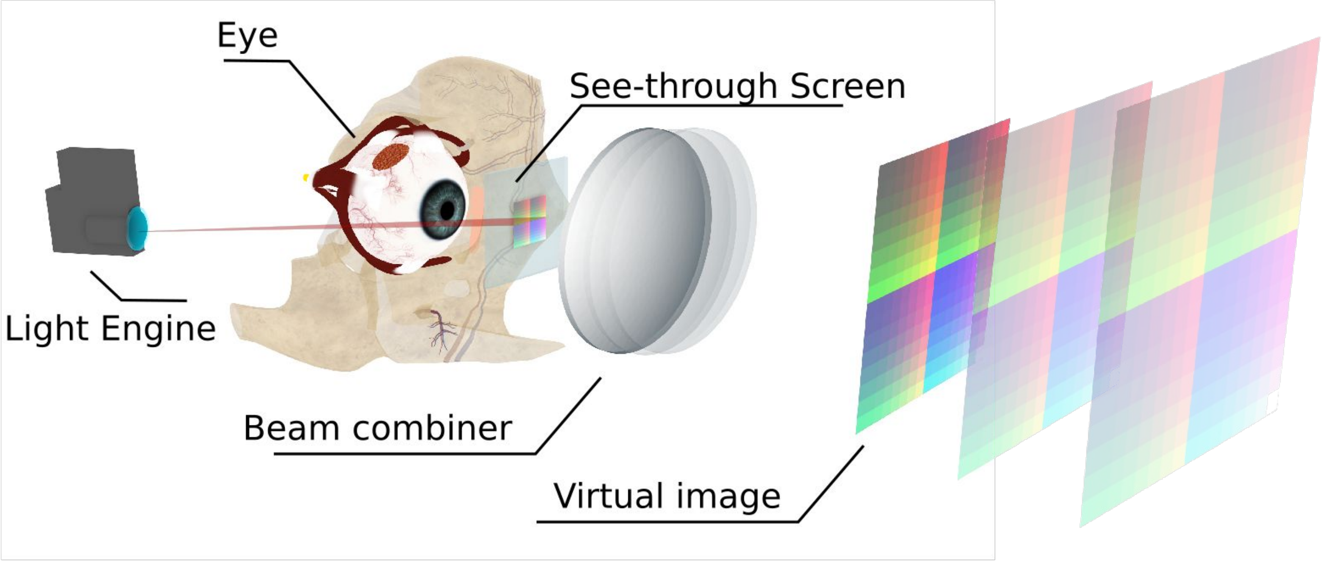

Eye
See-through Screen
Light Engine


Eye
See-through Screen
Light Engine


Eye
See-through Screen
Light Engine
Beam combiner


Eye
See-through Screen
Light Engine
Beam combiner
Virtual image

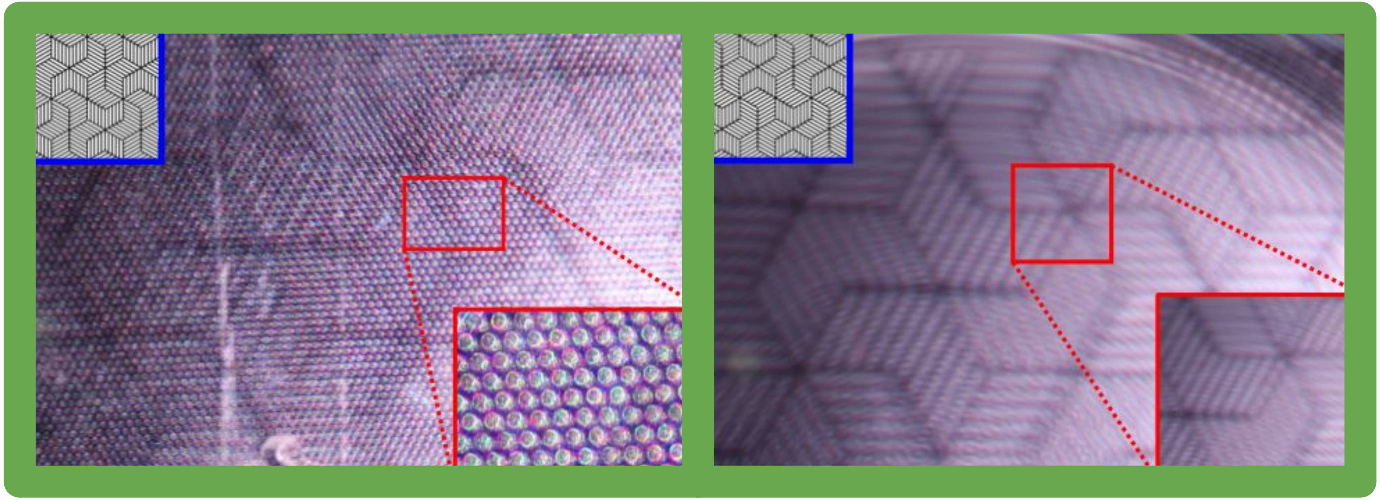

Eye
See-through Screen
Light Engine
Beam combiner
Virtual image


Eye
See-through Screen
Light Engine
Beam combiner
Virtual image

How to build it?

See-through Screens

Rotating diffusers



Cheap and dirty!

See-through Screens

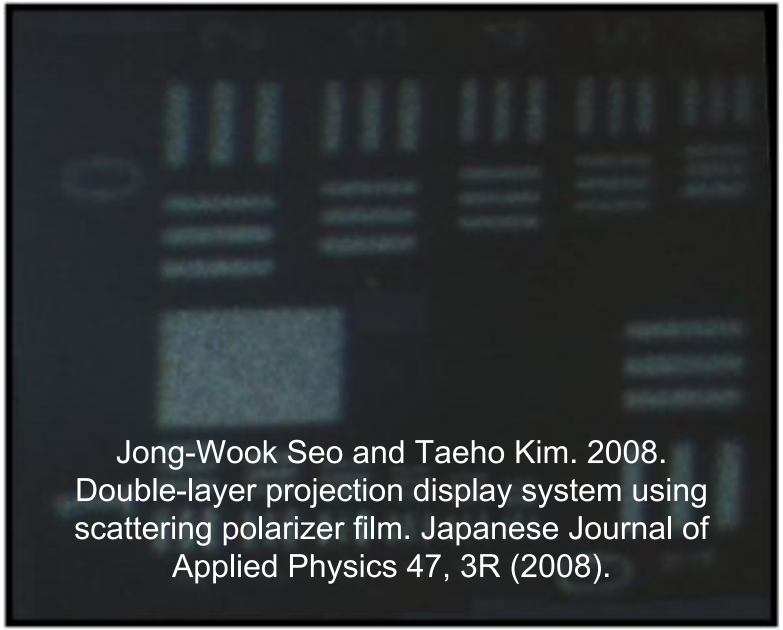
Rotating diffusers

Polarization
Selective Diffusers

Polarization Selective Diffuser

Samples from Nitto Japan

Limited screen size!



Jong-Wook Seo and Taeho Kim. 2008. Double-layer projection display system using scattering polarizer film. Japanese Journal of Applied Physics 47, 3R (2008).

See-through Screens

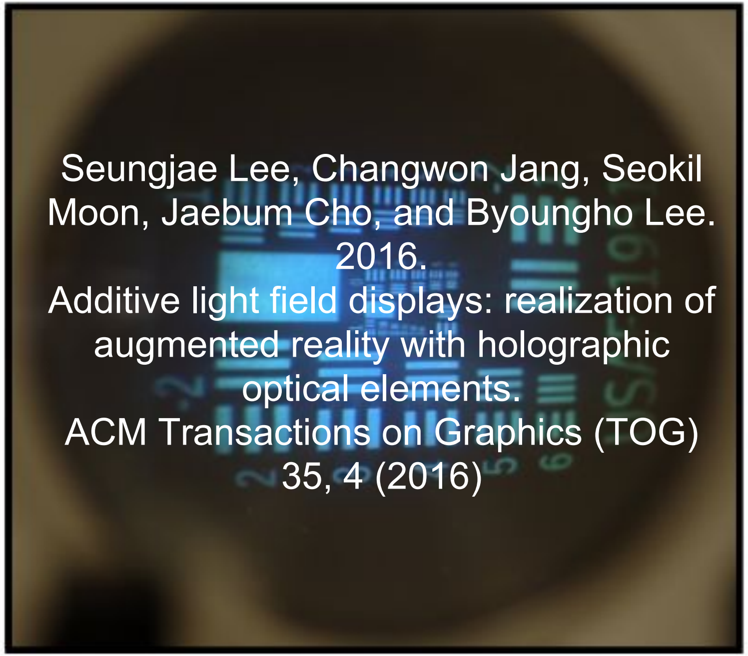
Rotating diffusers

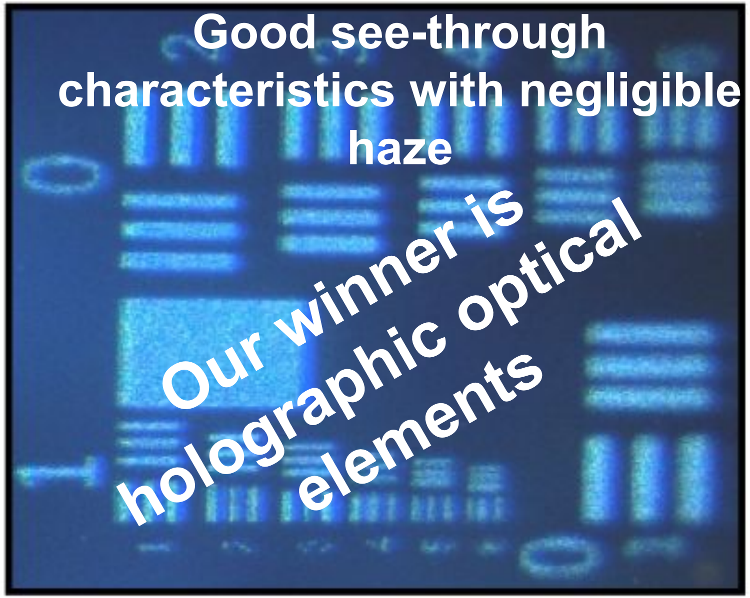
Polarization Selective Diffusers

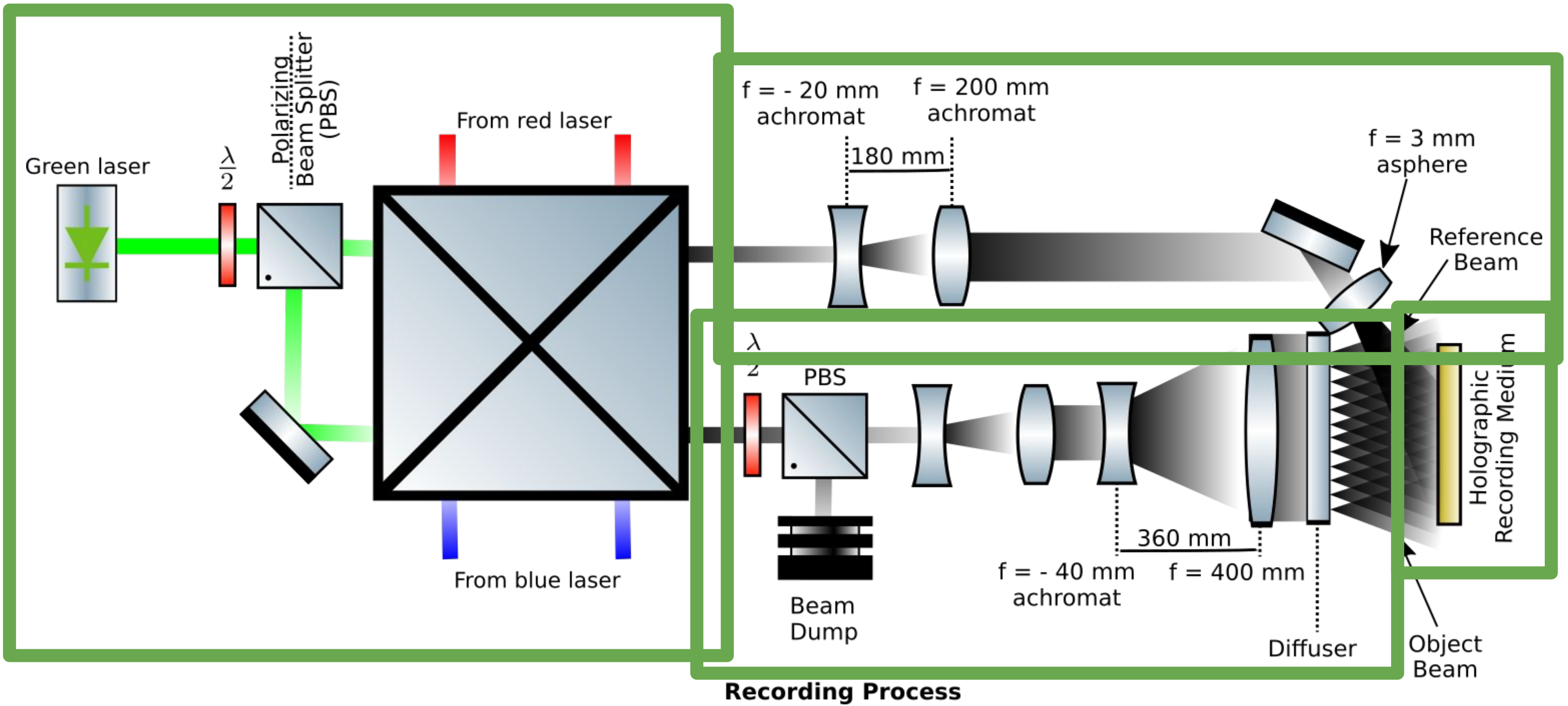
Holographic Optical Elements

In-house made

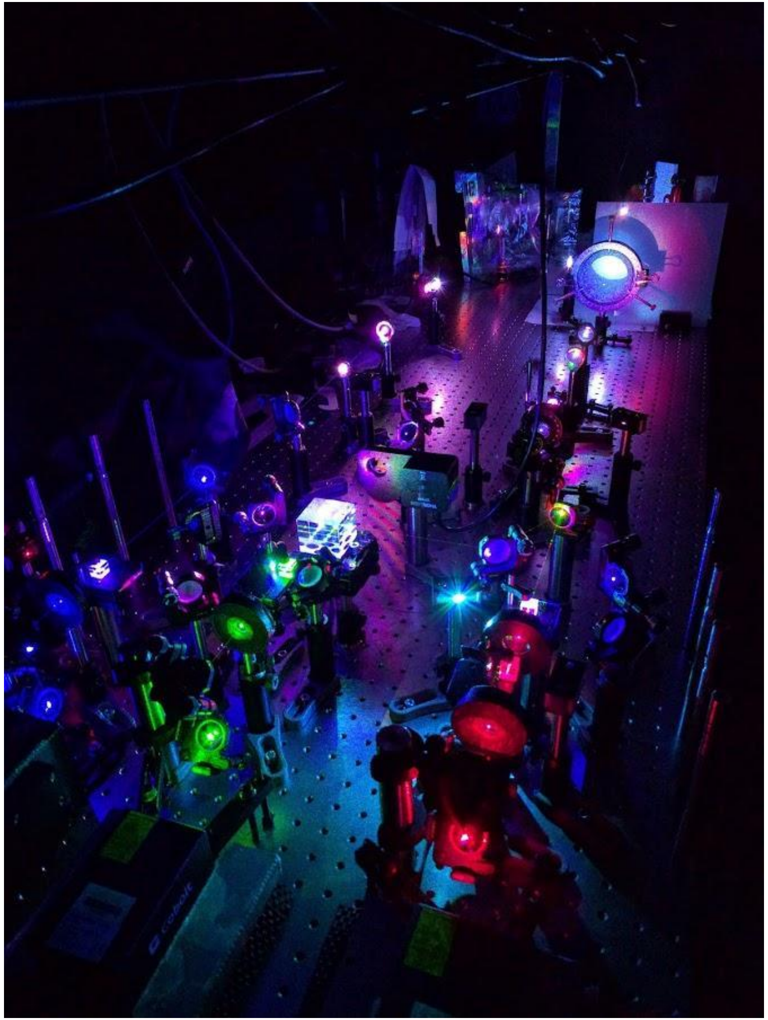
Holographic Optical Element





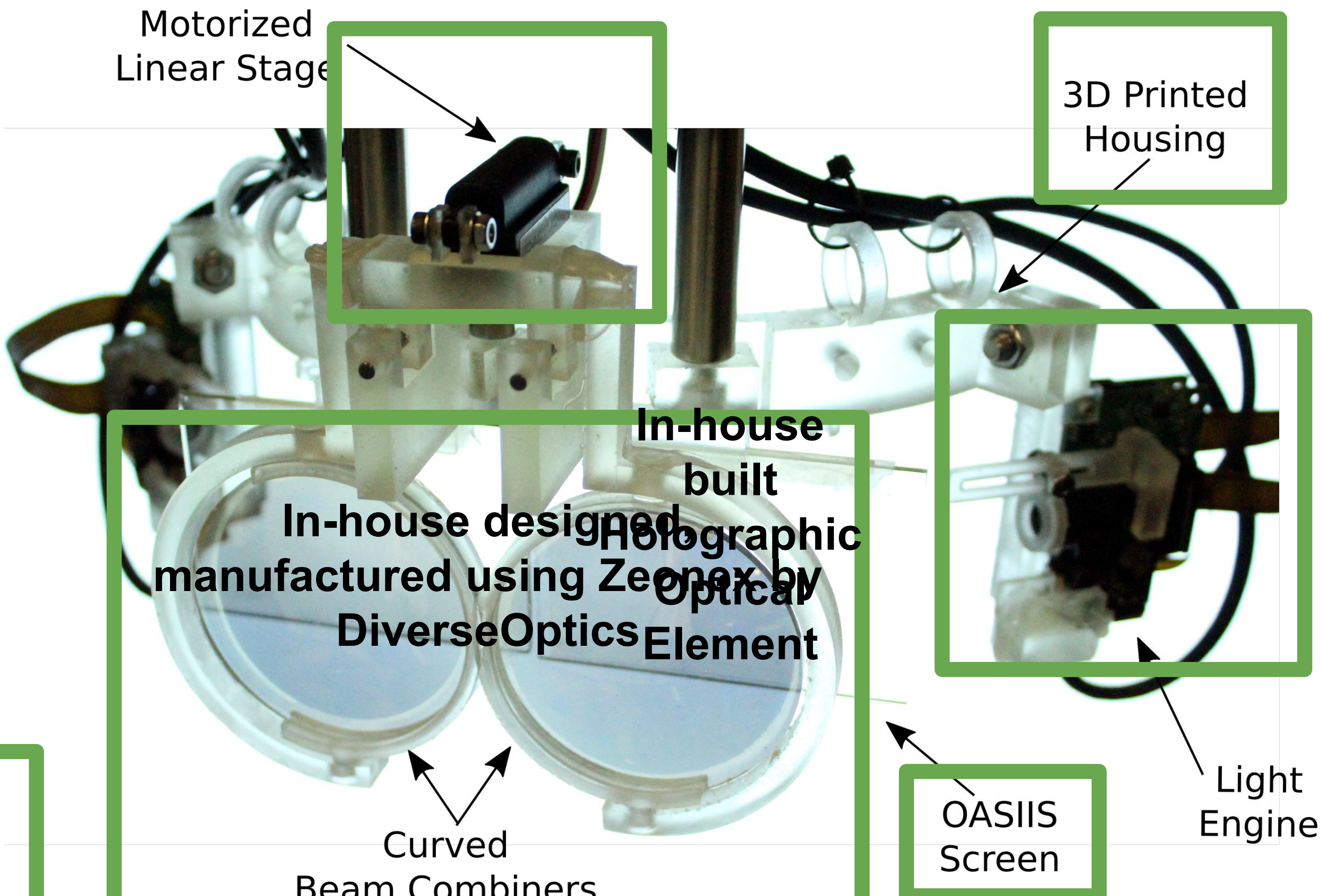


Note that this is an one time recording process, see-through screen are recorded to display dynamic content.



In-house analog holography setup

- Coherence length larger than 15 m, and 660-532-460 nm wavelengths for red, green, blue
- 120 grit ground glass diffuser from Edmund Optics
- Holographic recording medium from LitiHolo (16 um)



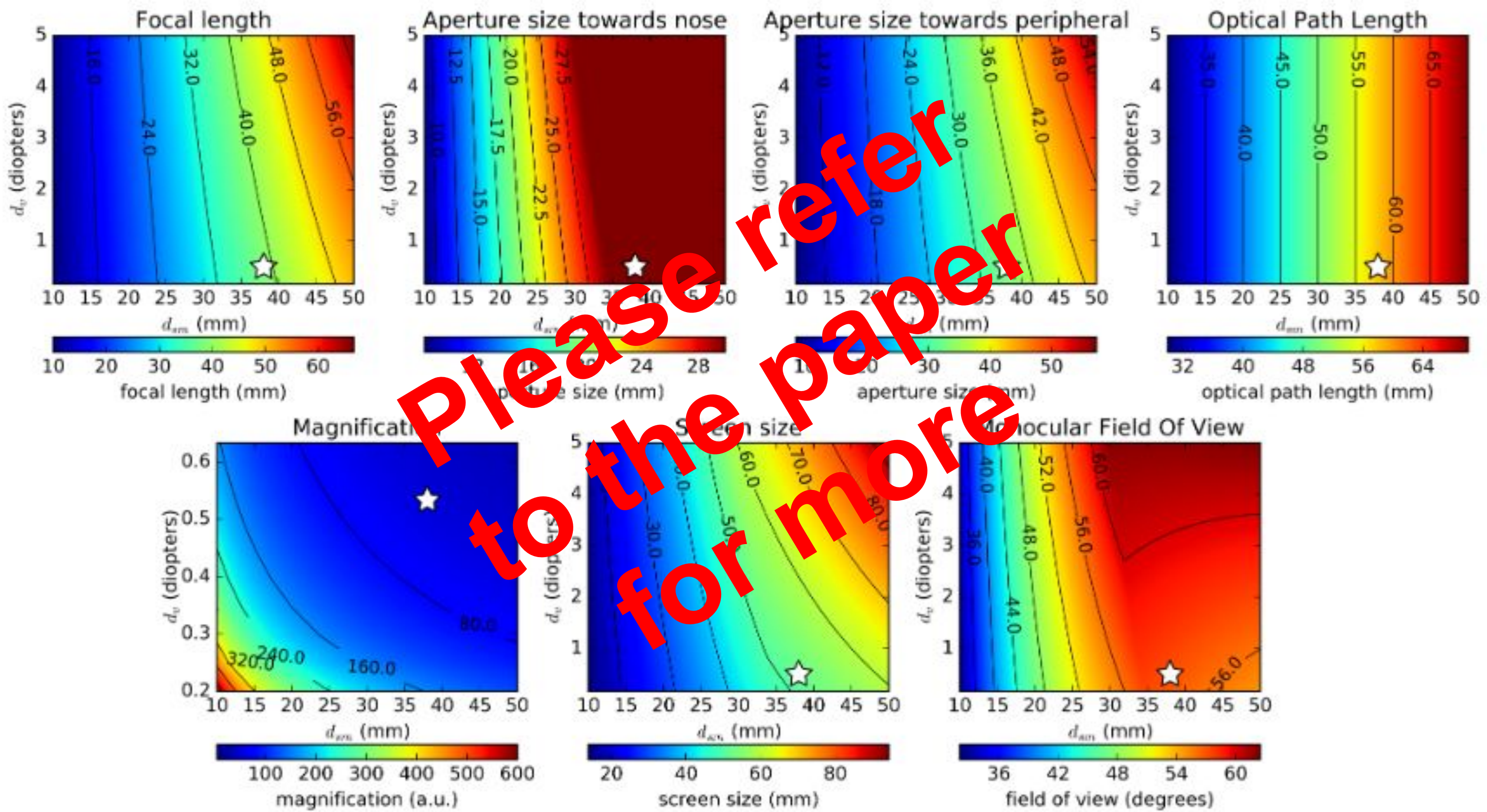
In-house designed, manufactured using Zeonex by DiverseOptics

In-house built Holographic Optical Element

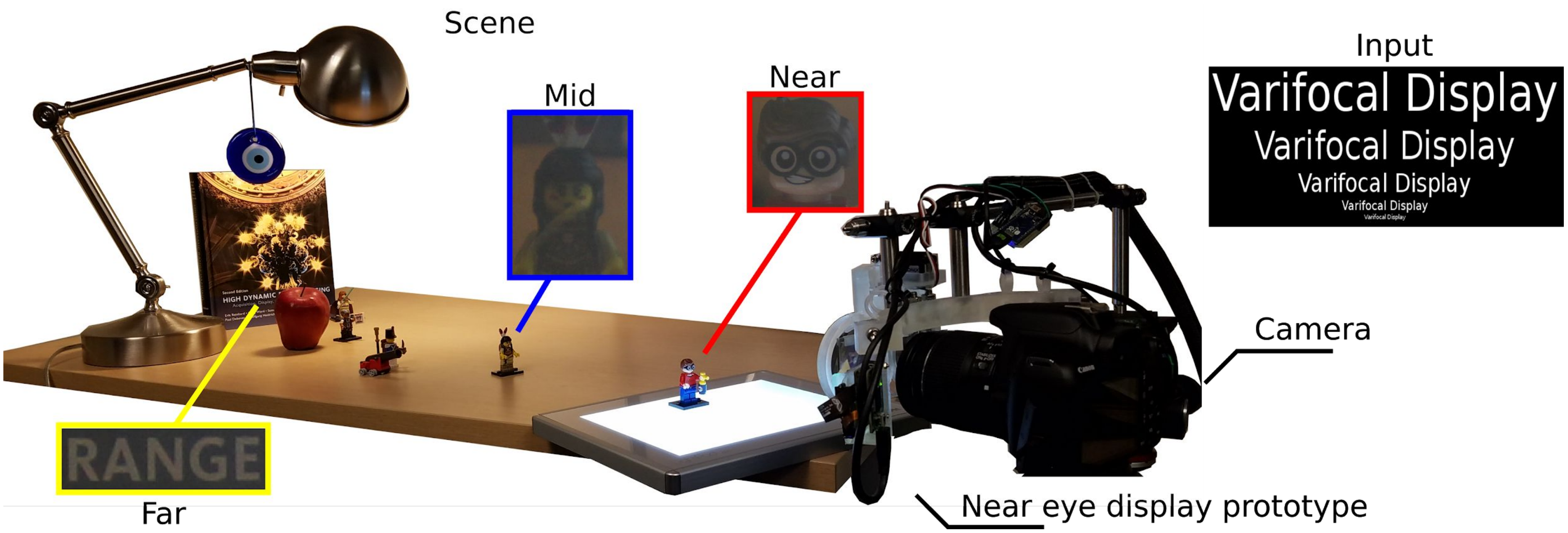
720p, 60 Hz, Liquid Crystal On Silicon (LCoS) from Imagine Optix

In-house OpenGL Based Renderer

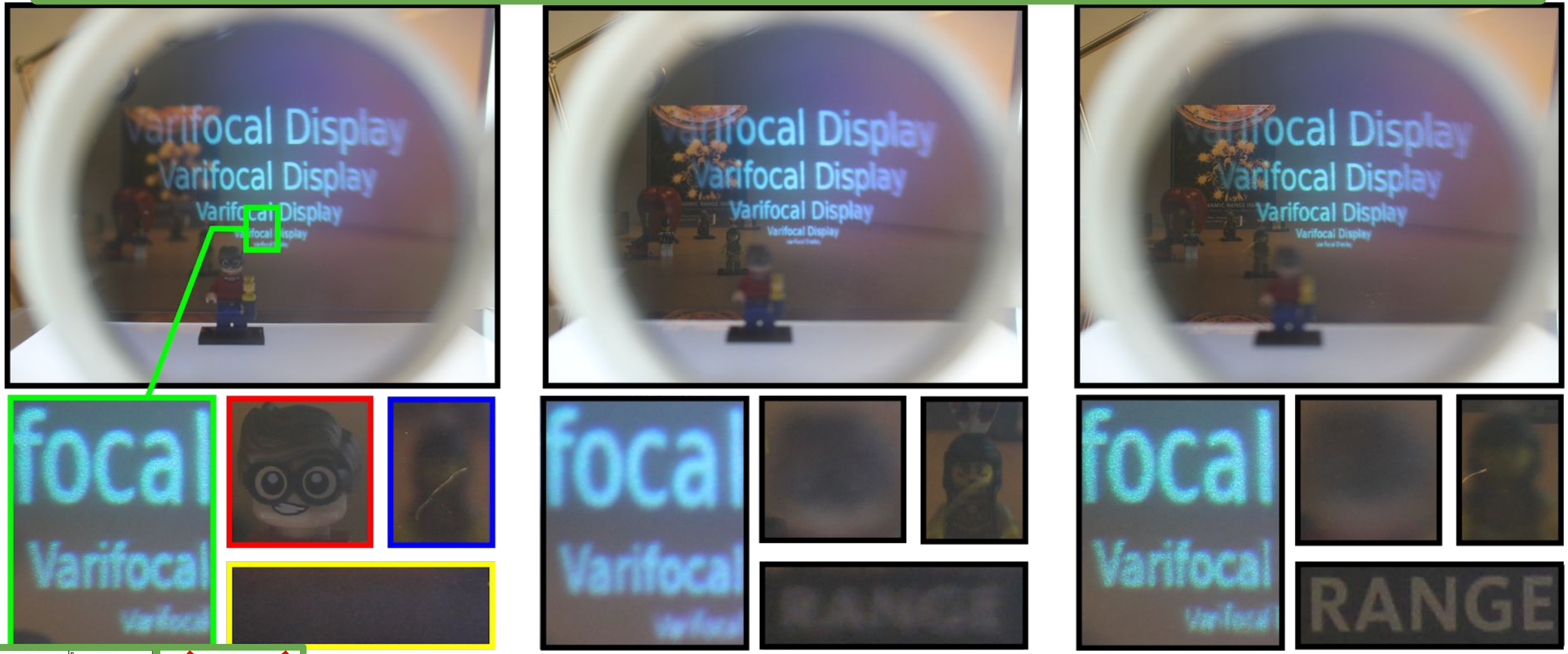
Wearable Varifocal Prototype

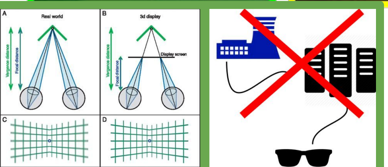

Focal length
Aperture size towards nose
Aperture size towards peripheral
Optical Path Length
d_v (diopters)
d_{sm} (mm)
focal length (mm)
aperture size (mm)
optical path length (mm)
Magnification
Screen size
Monocular Field Of View
magnification (a.u.)
screen size (mm)
field of view (degrees)
Please refer to the paper for more

Results

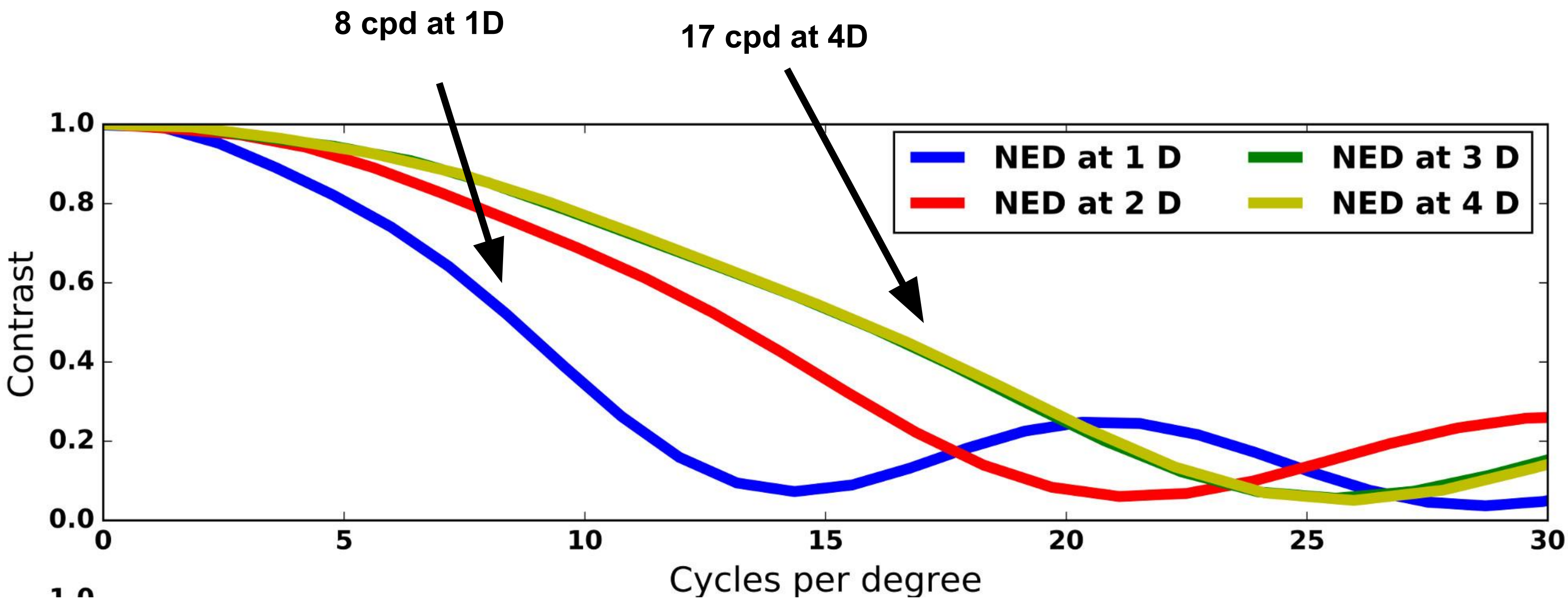

Scene
Mid
Near
Input
Varifocal Display
Varifocal Display
Varifocal Display
Varifocal Display
RANGE
Far
Camera
Near eye display prototype

Near	Mid	Far
25 cm	50 cm	100 cm

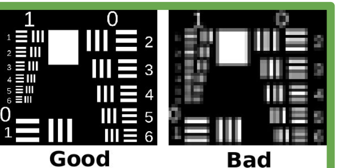


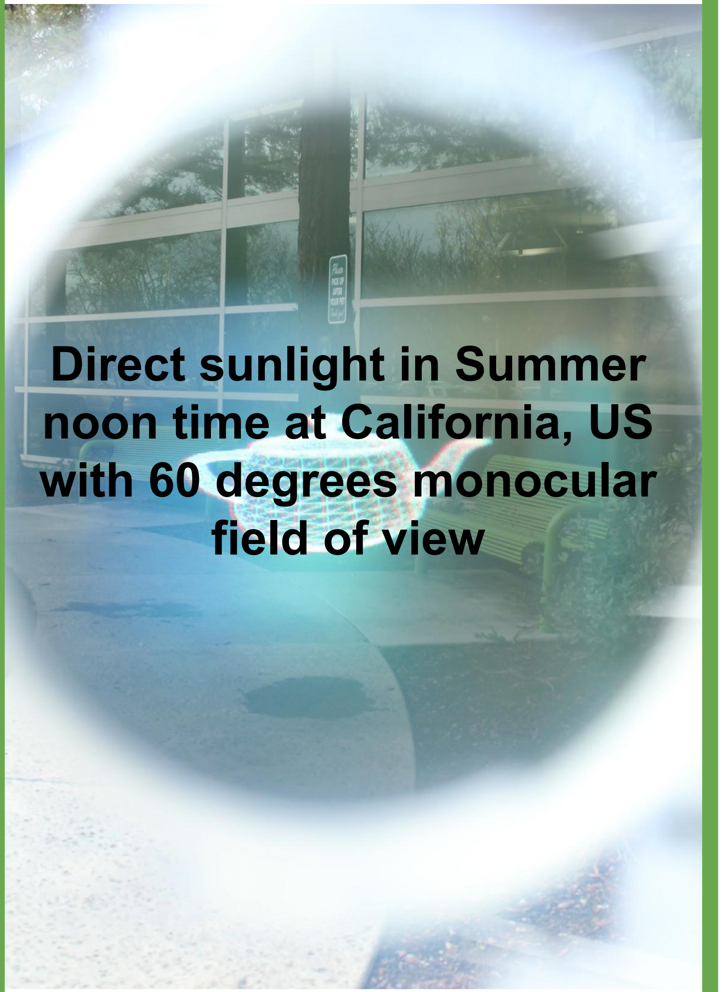


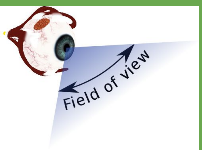
25 cm to infinity (6 m) with maximum 410 ms latency

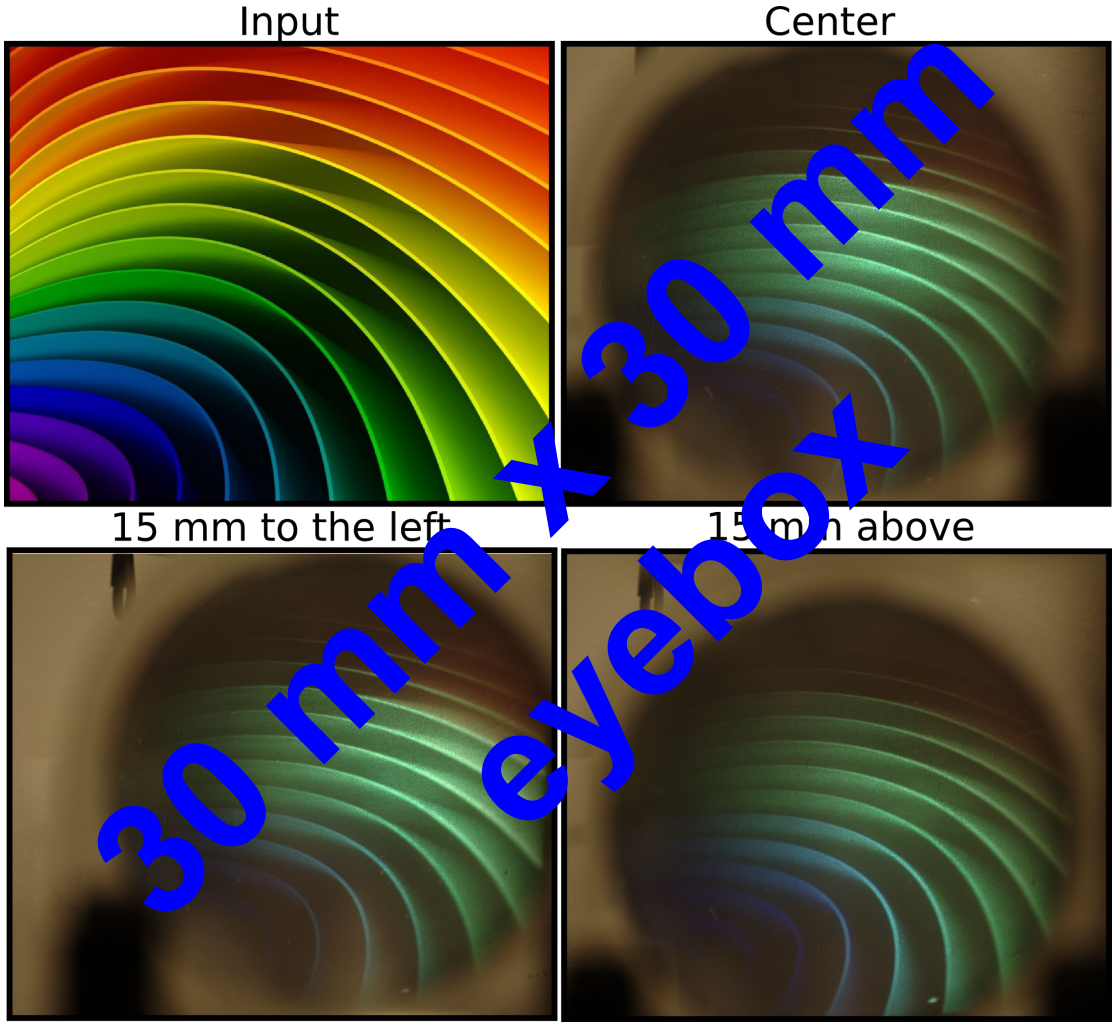


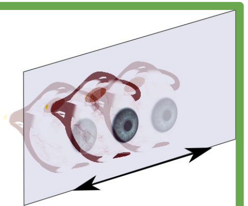
Peter D Burns. 2000. Slanted-edge MTF for digital camera and scanner analysis. Conference of Society for imaging science and technology, 135–138

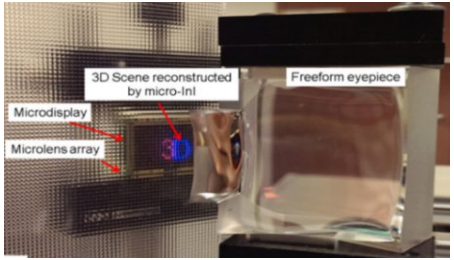



Direct sunlight in Summer noon time at California, US with 60 degrees monocular field of view


Field of view


Input
Center
15 mm to the left
15 mm above
30 mm x 30 mm
eyebox





Hong Hua and Bahram Javidi. 2014. A 3D integral imaging optical see-through head-mounted display. Optics express 22, 11 (2014).

Lightfield AR

No mechanically moving part or active [illegible]s, no need for a gaze tracker

Varifocal AR

Less compute demand, larger eyebox, better resolution, and much wider field of view

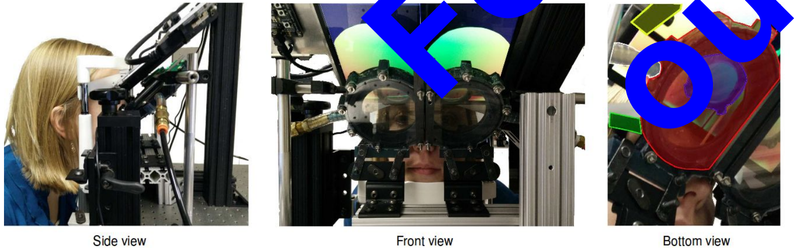
Andrew Maimone, Andreas Georgiou, and Joel Kollin. 2017. Holographic Near-Eye Displays for Virtual and Augmented Reality. ACM Transactions on Graphics 36 (2017).

Holography AR

No mechanically moving part or active parts, better form-factor

Varifocal AR

Much less compute demand, much larger eyebox,



Dunn, David, et al. "Wide Field Of View Varifocal Near-Eye Display Using See-Through Deformable Membrane Mirrors." *IEEE Transactions on Visualization and Computer Graphics* 23.4 (2017): 1322-1331.

Varifocal AR

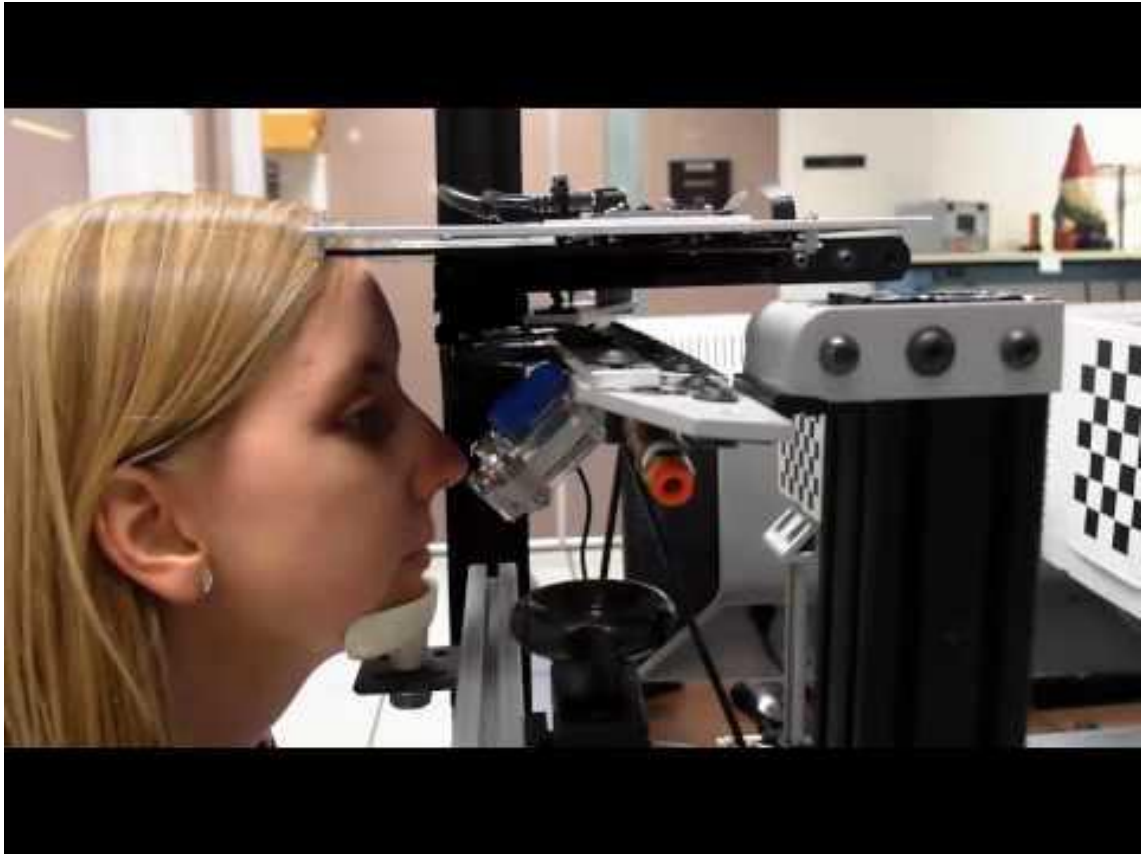
Much faster focus change

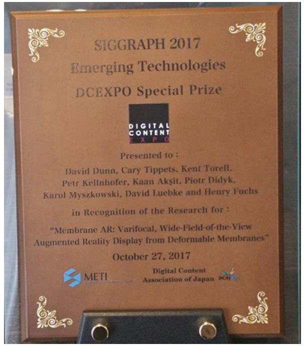
Varifocal AR

Much better form factor, much larger eyebox

For more see our paper

Varifocal display proposal II



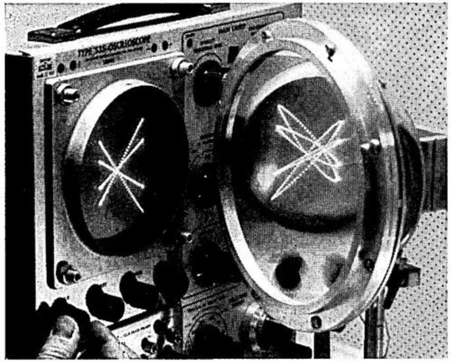


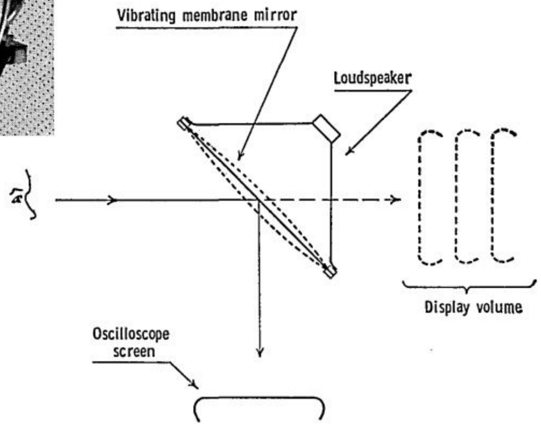
SIGGRAPH 2017 DCEXPO SPECIAL PRIZE!

David Dunn, Cary Tippets, Kent Torell, Petr Kellnhofer, Kaan Akşit, Piotr Didyk, Karol Myszkowski, David Luebke, and Henry Fuchs. “Wide Field Of View Varifocal Near-Eye Display Using See-Through Deformable Membrane Mirrors.” *IEEE Transactions on Visualization and Computer Graphics* 23, no. 4 (2017)

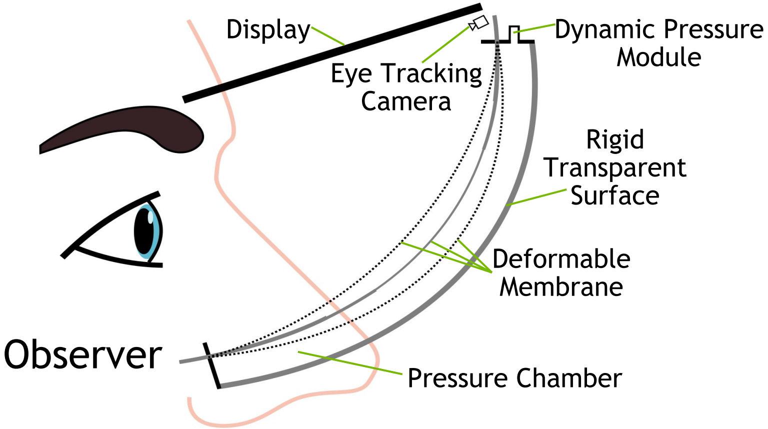
VOLUMETRIC DISPLAYS

- Vibrating membrane mirror
- Refresh dictated by speed of display/depth resolution
- Defined volumetric range
- Small diagonal FOV
- Not see-through





- Dynamic focal depth
- Wide field of view
- Single element optics



Membrane

Dynamic Pressure System

Membrane Tracking System

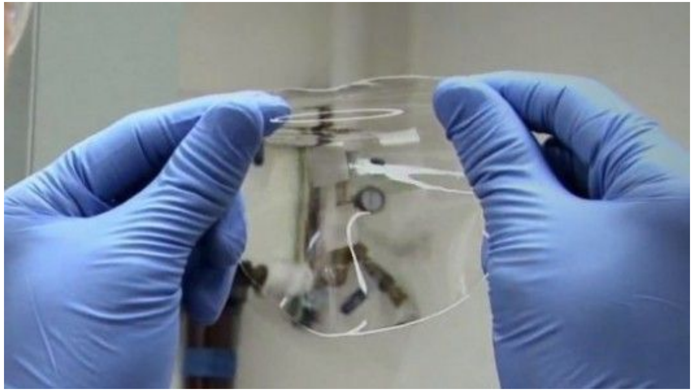
Eye Tracking System

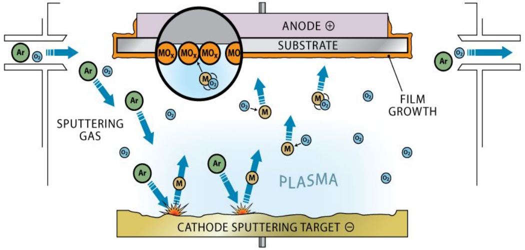
How to build it?

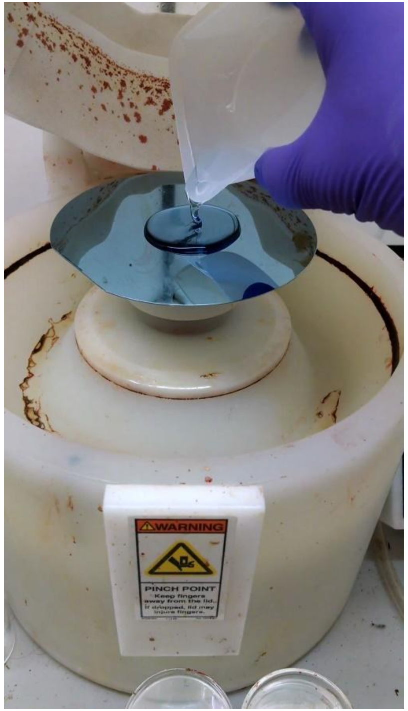
Membrane Creation: Material

Polydimethylsiloxane [PDMS]

- Silicon-based organic polymer
- Optically clear
- Viscoelastic material
- Sputter coated with silver to enhance reflection

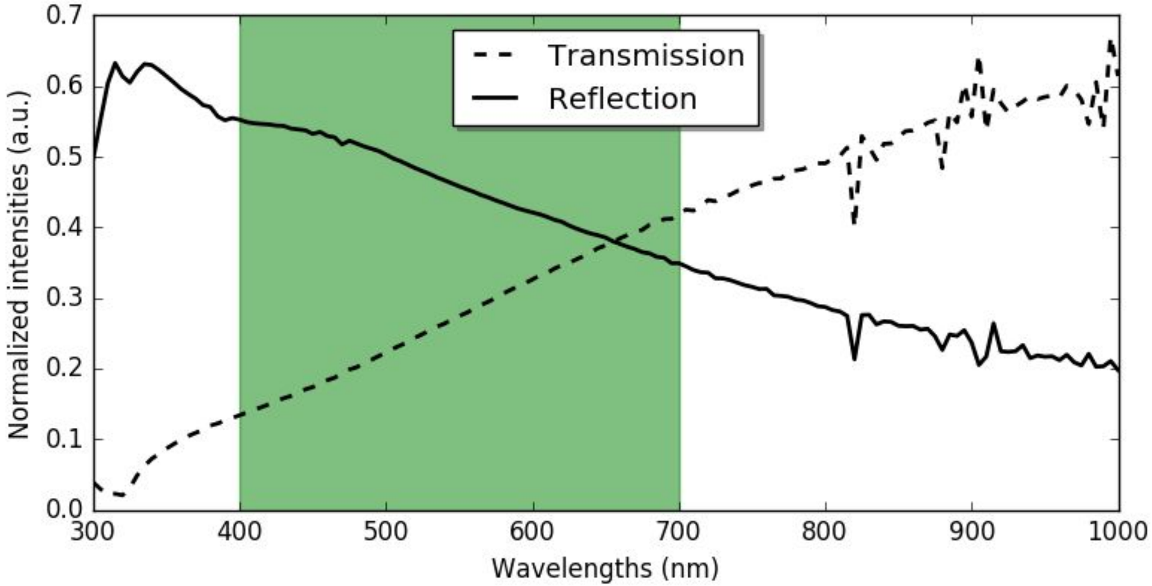




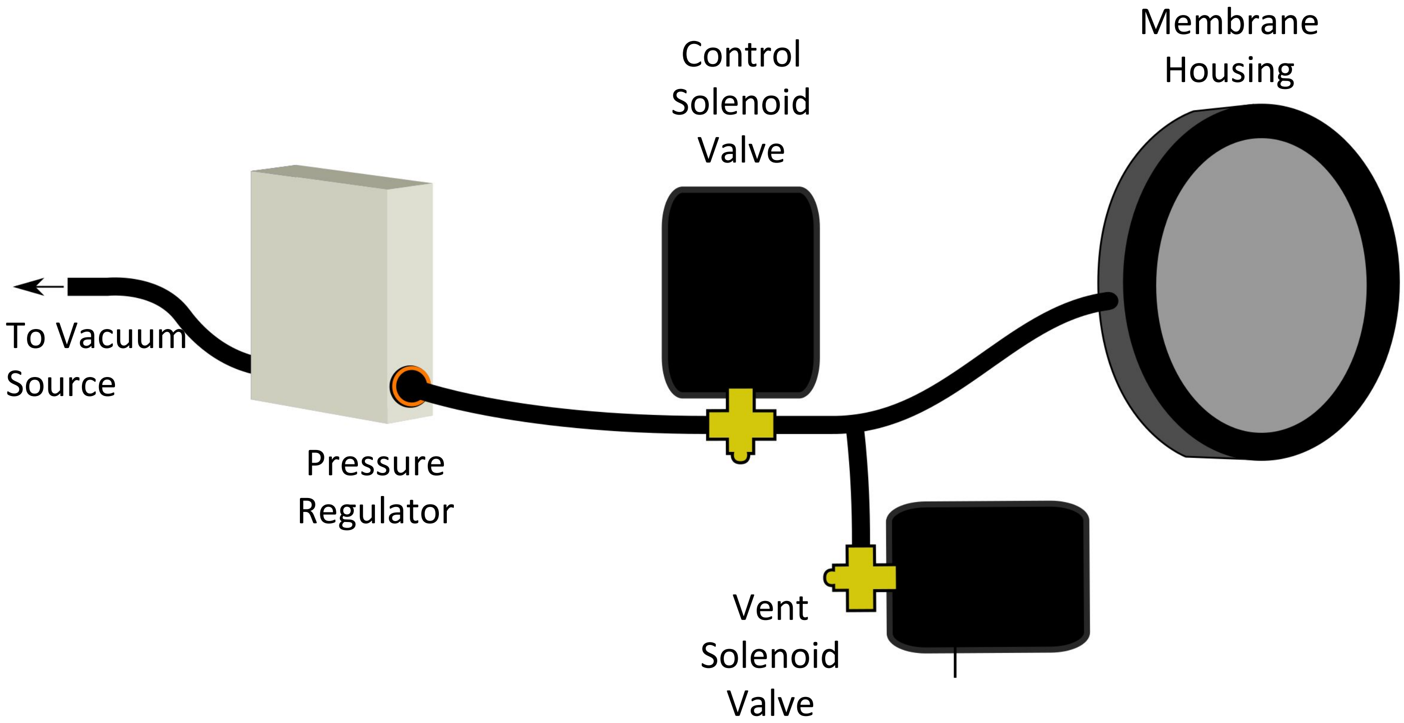


https://www.youtube.com/watch?v=5boywxr8ot4
http://clearmetalsinc.com/technology/

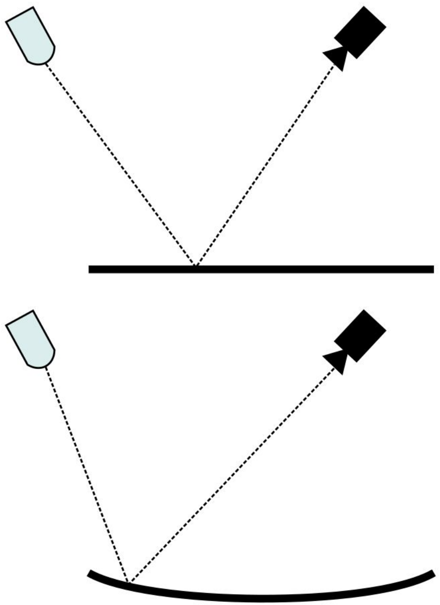
Reflection is Wavelength Dependent



Vacuum System



LED Camera System



Feedback to know the shape of the membrane

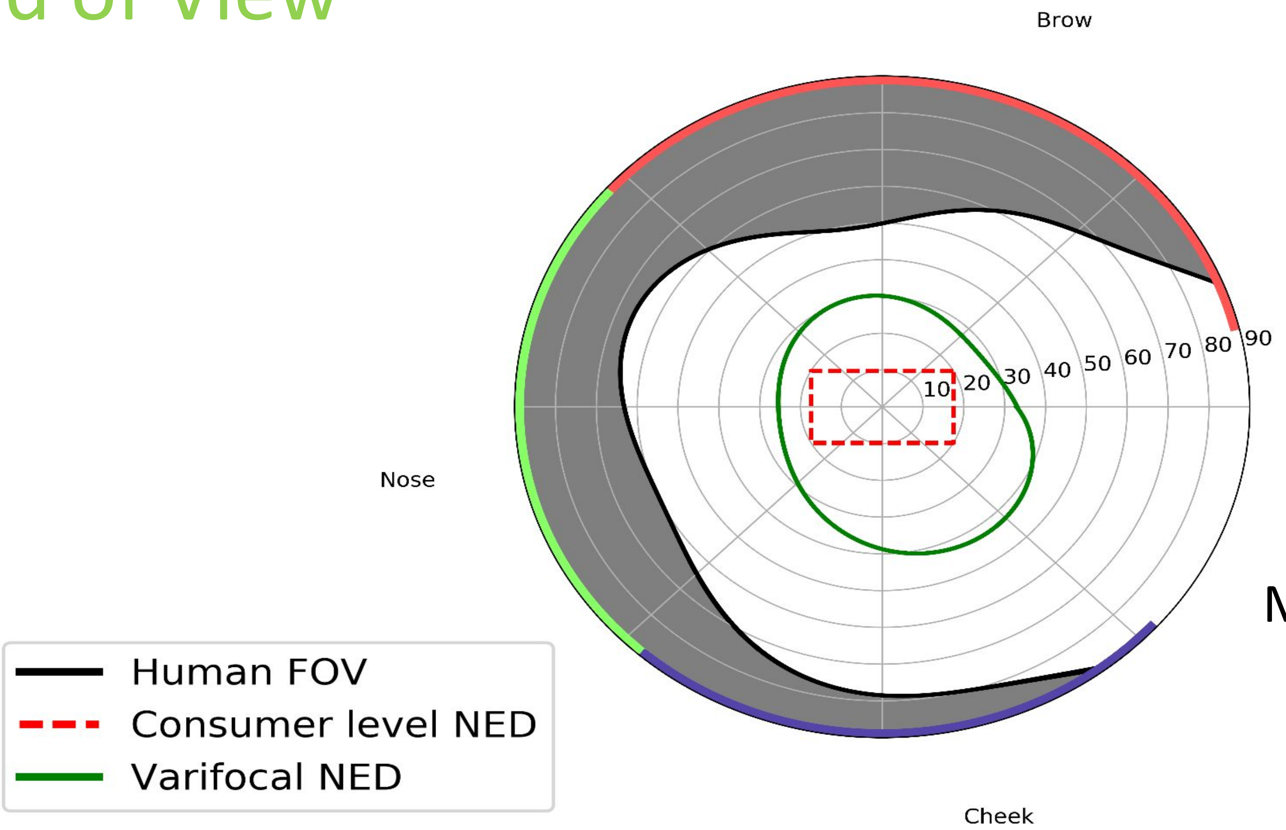
As the membrane deforms the LED's reflection moves

Blob detection is used to locate and track the motion

Uses infrared light to not distract the user

Results

Field of View



Monocular FOV:

75° diagonal

55° horizontal

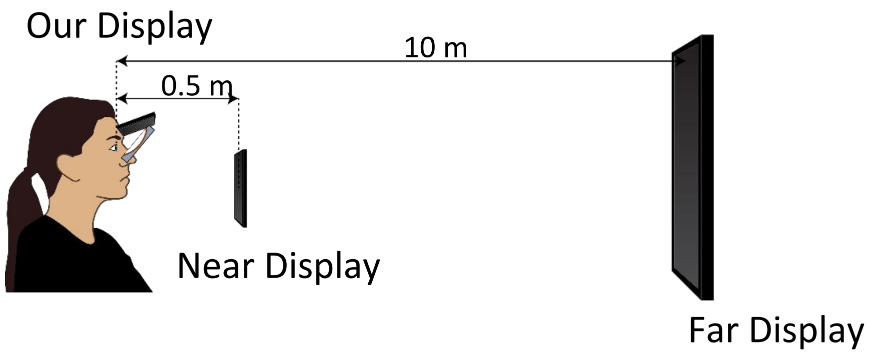
65° vertical

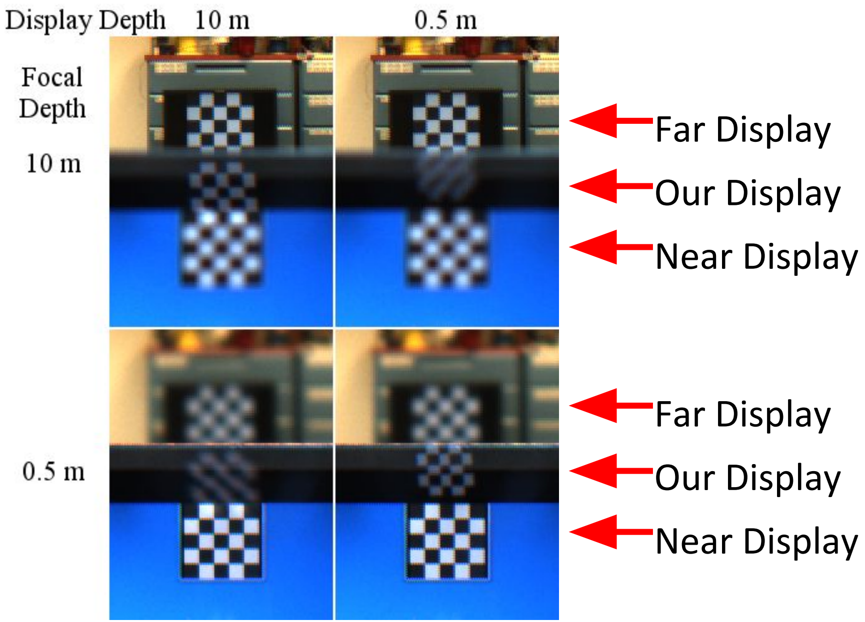
Focal Depth

7 diopter range (15cm - infinity)

Under 300ms from far to near

Under 300ms from near to far





Focus Consistency

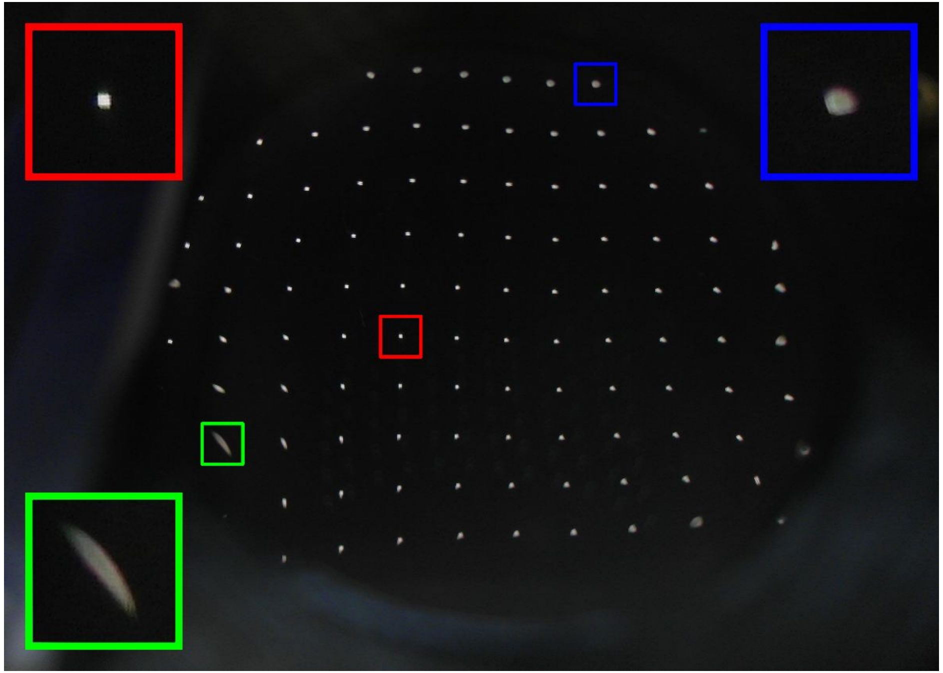
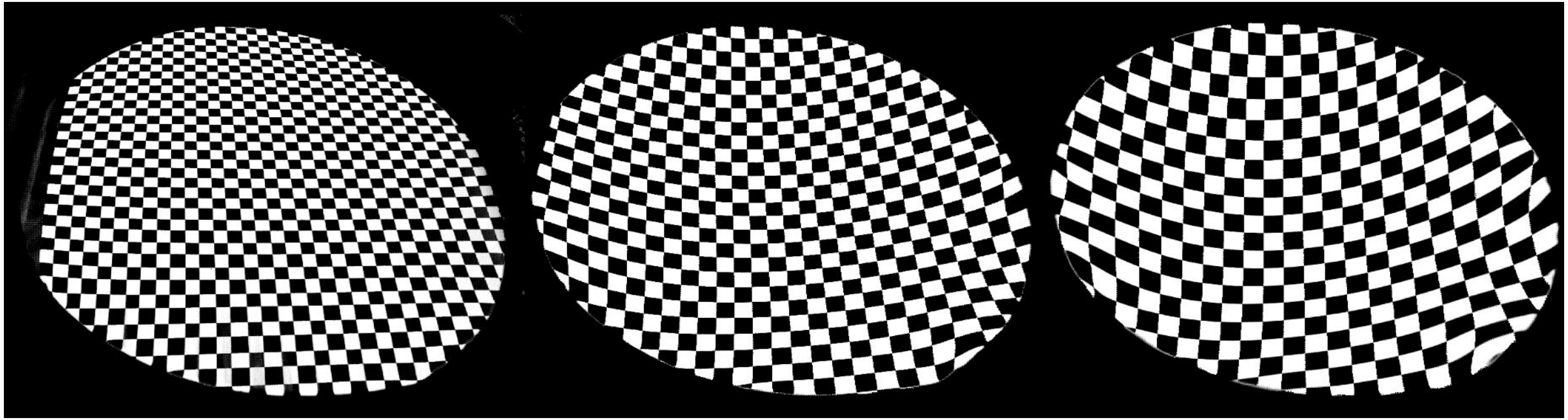
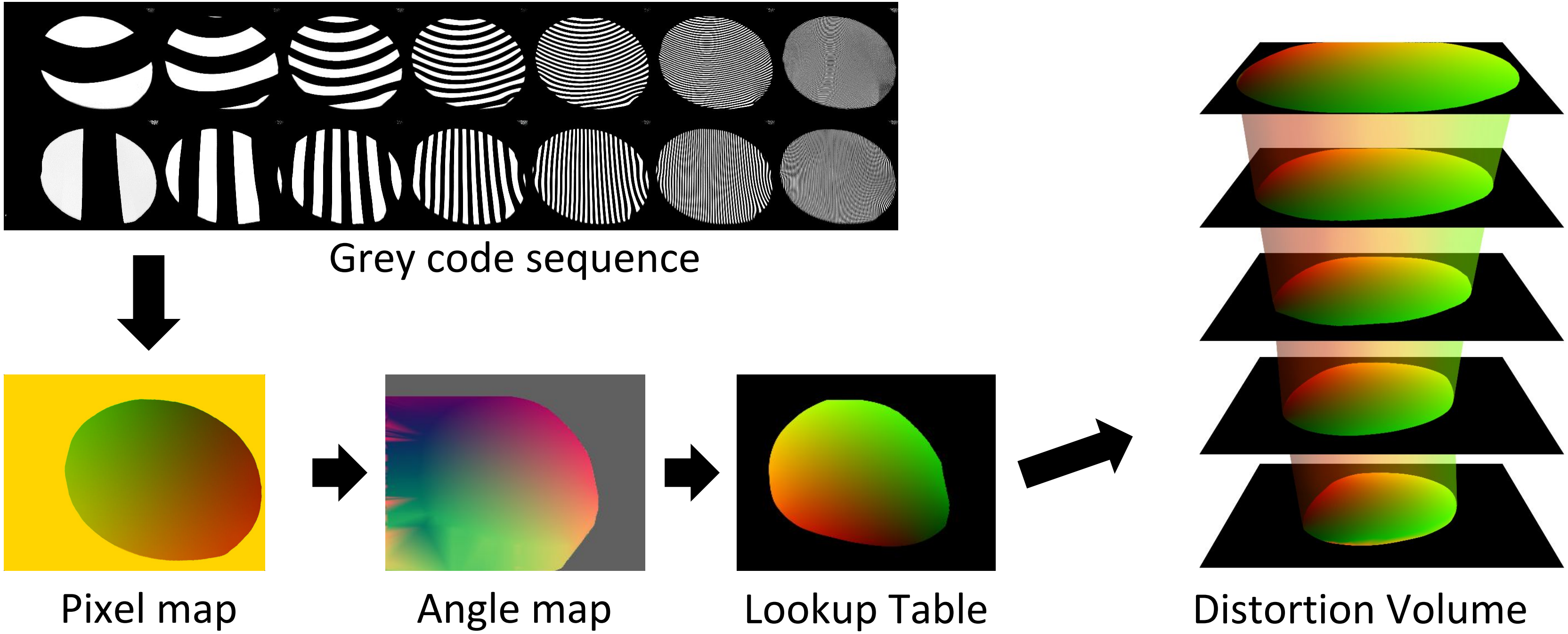


Image Distortion

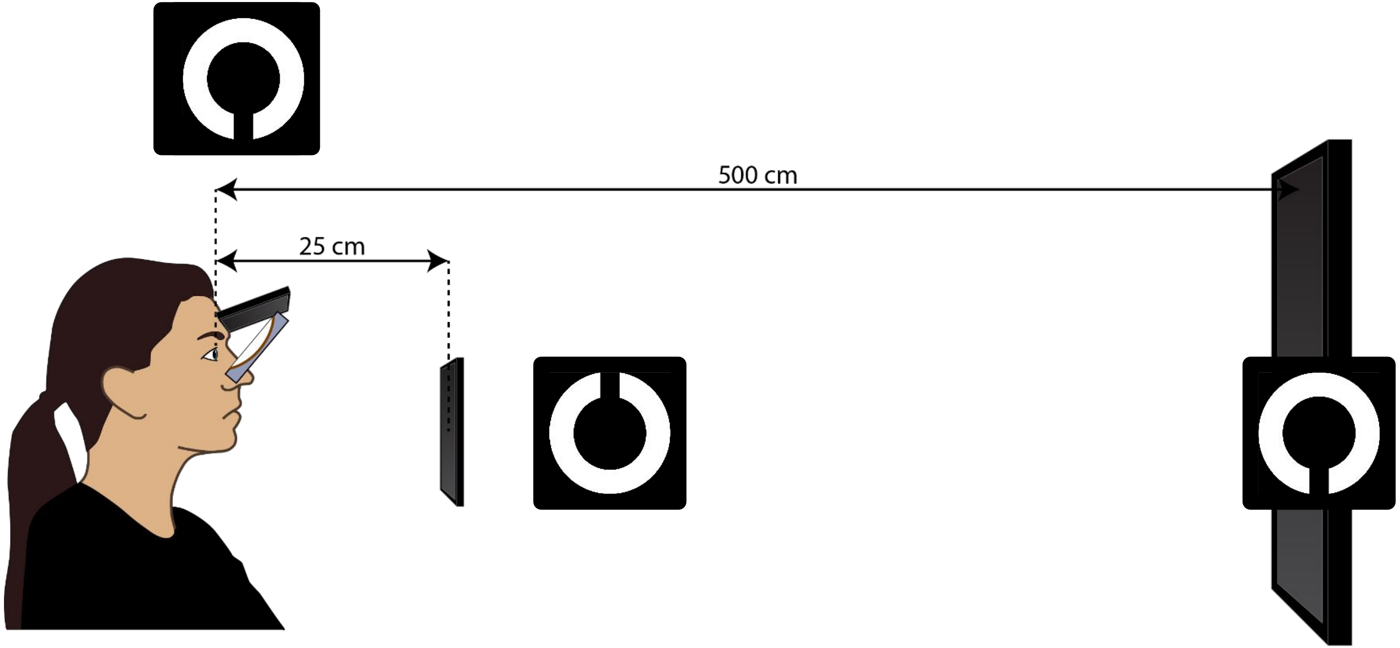


Near Mid Far

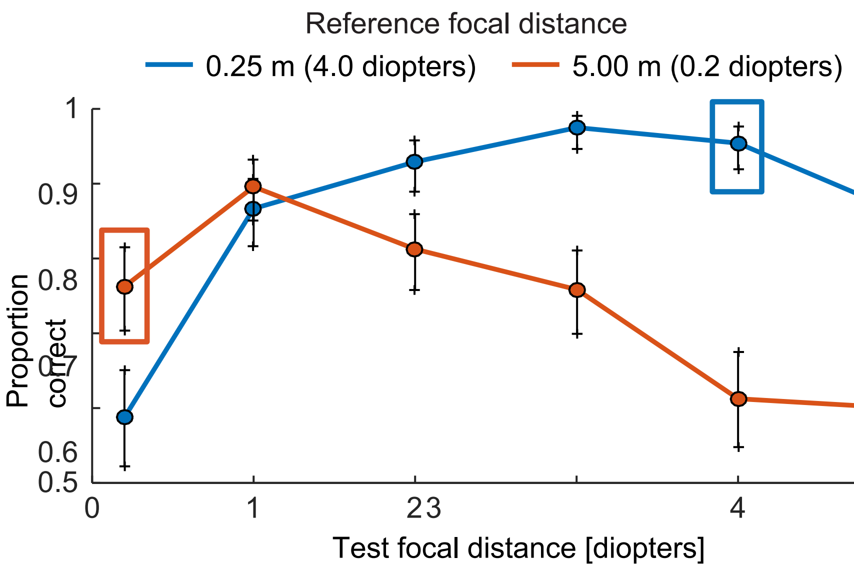
Distortion Correction



Perceptual Experiment



Perceptual Experiment



3D printing optics

Formlabs 2	Formmech 508DT	Norland Optical Adhesive	Clear Acrylic
Price: 4999 USD	Price: 7413 USD	Price: 30 USD	Price: 10 USD

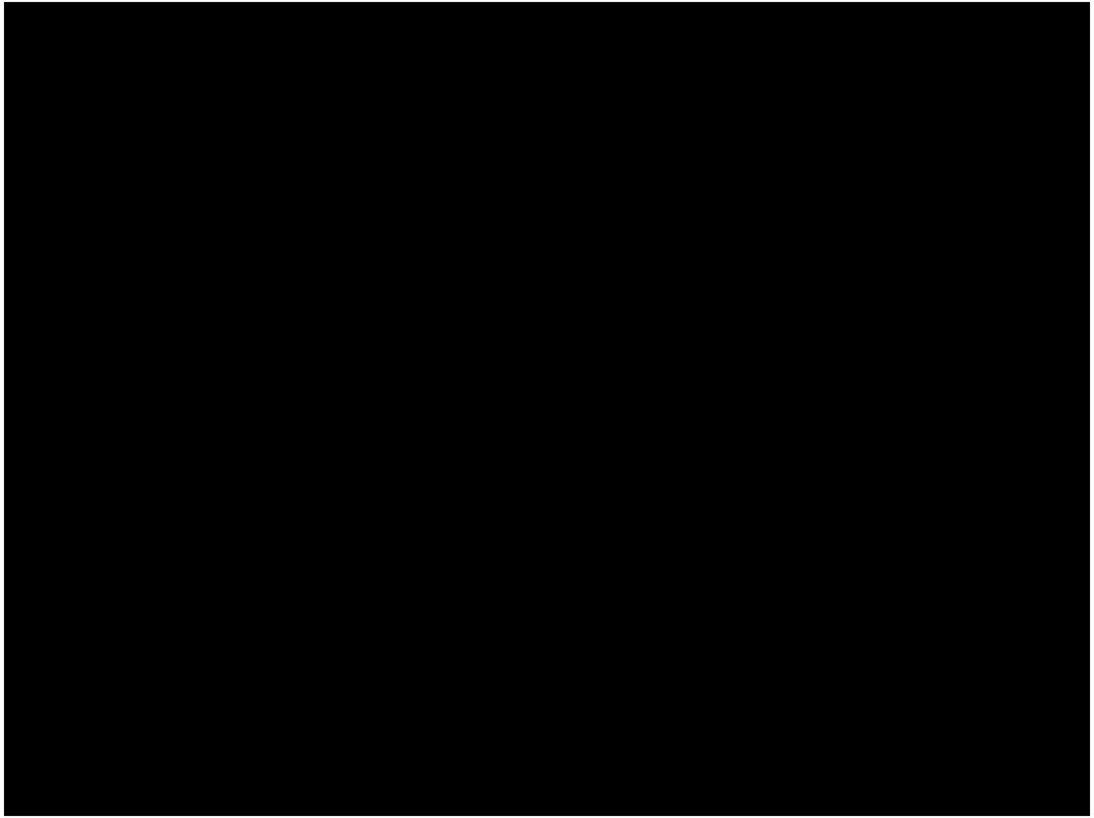
Investment : ~15-20k USD + you + short processing times (1 day)

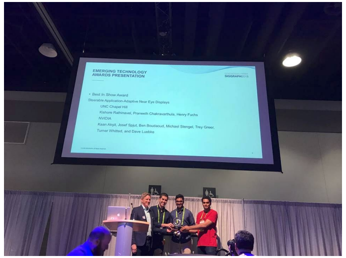
---> Good for fast prototyping <---

Manufacturing Application-Driv
Near-Eye Displays
NVIDIA
Manufacturing freeform optical components is a slow, expensive, and labor-intensive task that restricts computation
optical designers to only existing components which may be sub-optimal, and limits their ability to iterate and explore new
typical required investment in the orders of 1-5 Million USD
trained personnel is required, labor intesive task, and slow
Optical adhesive Cost: < 200 USD
A series of photographs showing optical components built using our manufacturing technique based on 3D printing, optical bonding, and vacuum forming
Two photographs showing a view of a mobile phone display magnified by (Left) an off-the-shelf lens (Thorlabs LA1401), and (right) a clone of the same lens manufactured using our technique, to provide a visual comparison of optical quality
hoose the optically challenging scenario of near-eye displays to evaluate both our manufacturing
, and our computational methodology for calculating interchangeable freeform surfaces for a given
lay optical layout. We design and manufacture a completely untethered near-eye display prototype.

Op
Gl ed
ba
Le

Printed Near-Eye Displays





SIGGRAPH 2018
BEST IN SHOW
AWARD

Kaan Akşit, Praneeth Chakravarthula, Kishore Rathinavel, Youngmo Jeong, Rachel Albert, Henry Fuchs and David Luebke. “Manufacturing Application-Driven Foveated Near-eye Displays” *(SUBMITTED FOR REVIEW)* (2018)

What is next?

“The Last Slide”

New layouts based on novel see-through screens enables on-axis/off-axis paths: better resolution, field of view and eyebox!

More resolutions, more field of view, slimmer form factor?

Merging with others?

Prime time proof for varifocal?

Thank you for listening

Nvidia Research
http://research.nvidia.com

Kaan Akşit,
kaksit@nvidia.com
https://kaanaksit.com

SIGGRAPH ASIA 2018 TOKYO

Rafał K. Mantiuk

HDR, displays & low-level vision

SIGGRAPH Asia Course on Cutting-Edge VR/AR Display Technologies

UNIVERSITY OF CAMBRIDGE

HDR & VR ?

- Do we have HDR VR headsets?

http://www.oculusvr.com/

- OLED contrast 1,000,000:1

ToC

- HDR in a nutshell
- Display technologies in VR
- Perception & image quality
- Example: Temporal Resolution Multiplexing

Dynamic range

Luminance

$$\frac{\max L}{\min L}$$

(for SNR>3)

Dynamic range (contrast)

- As ratio:

$$C = \frac{L_{\max}}{L_{\min}}$$

  - Usually written as C:1, for example 1000:1.

- As "orders of magnitude" or log10 units:

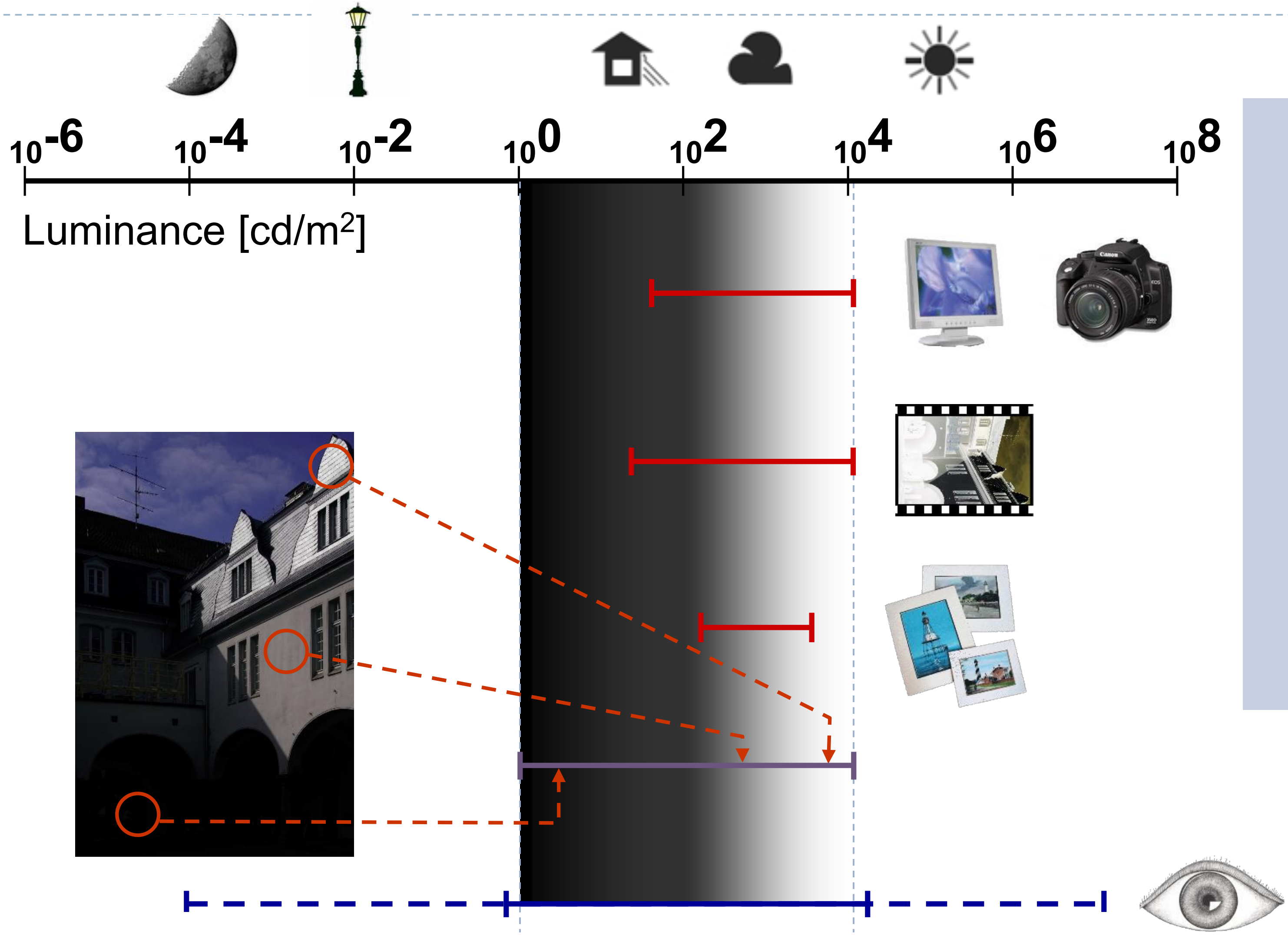
$$C_{10} = \log_{10} \frac{L_{\max}}{L_{\min}}$$

- As stops:

$$C_2 = \log_2 \frac{L_{\max}}{L_{\min}}$$

One stop is doubling
of halving the amount of light

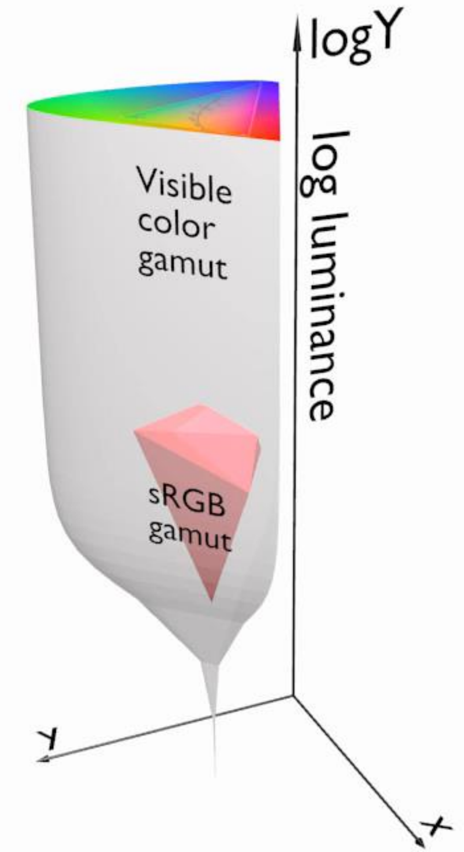
High dynamic range (HDR)



Dynamic Range
1000:1
1500:1
30:1

Visible colour gamut

- The eye can perceive more colours and brightness levels than
  - a display can produce
  - a JPEG file can store
- The premise of HDR:
  - Visual perception and not the technology should define accuracy and the range of colours
  - The current standards not fully follow to this principle



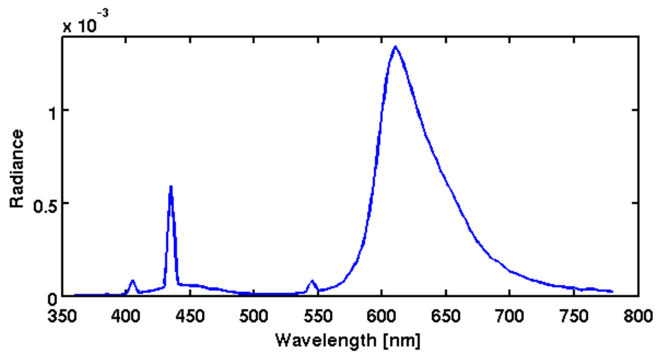
Luminance

- Luminance – how bright the surface will appear regardless of its colour. Units: cd/m^2

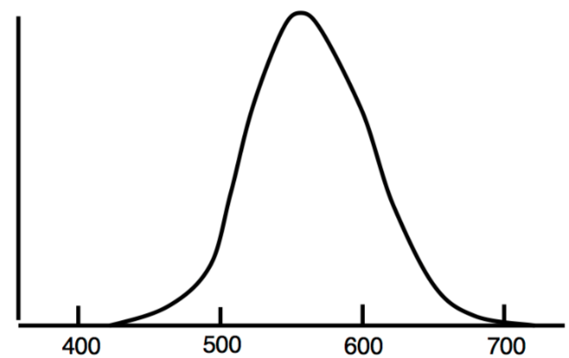
Luminance

$$L_V = \int_{350}^{700} kL(\lambda)V(\lambda)d\lambda \qquad k = \frac{1}{683.002}$$

Light spectrum (radiance)



Luminous efficiency function (weighting)



Luminance and Luma

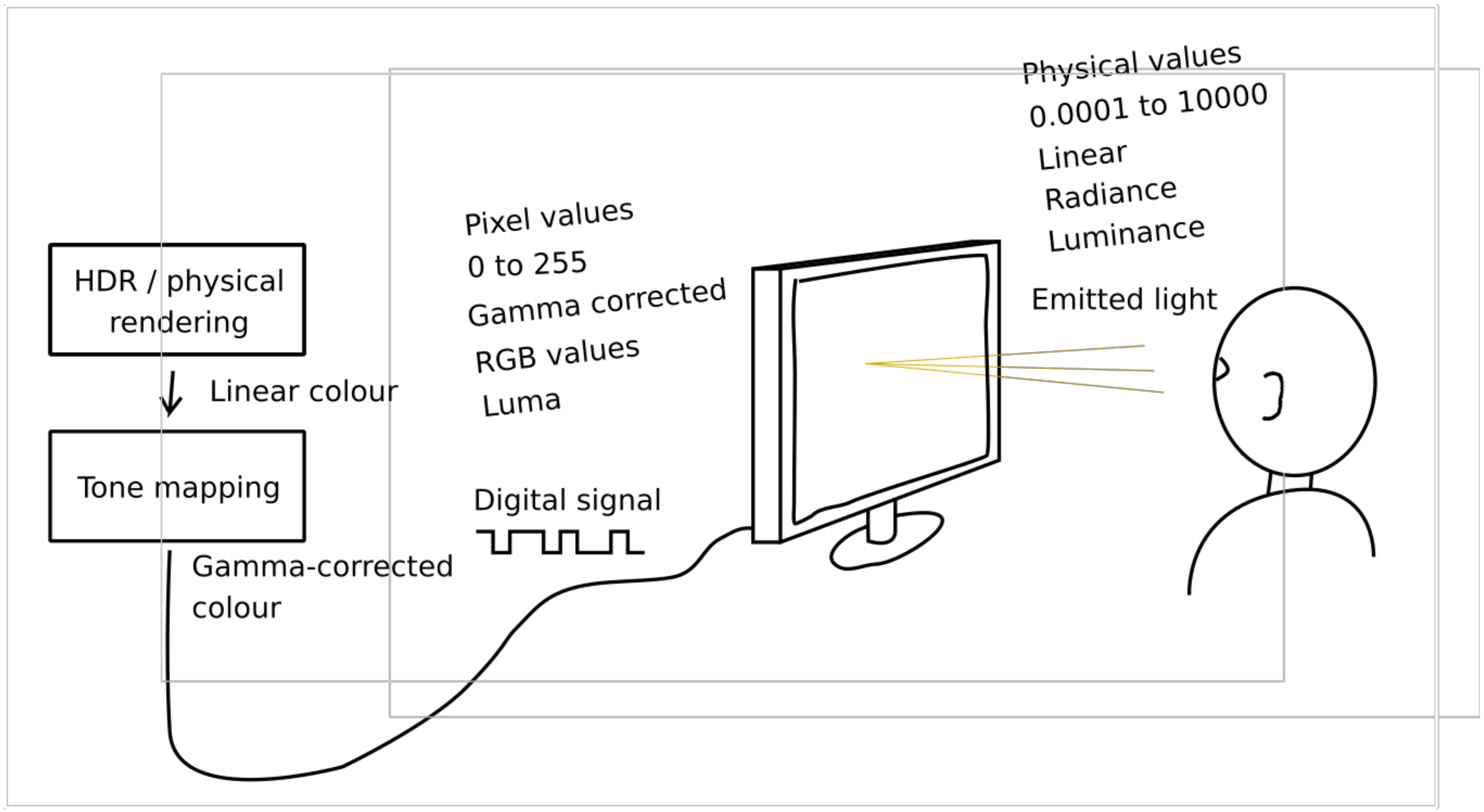
- Luminance
  - Photometric quantity defined by the spectral luminous efficiency function
- $L \approx 0.2126\,R + 0.7152\,G + 0.0722\,B$
- Units: cd/m^2

- Luma
  - Gray-scale value computed from LDR (gamma corrected) image
- $Y = 0.2126\,R' + 0.7152\,G' + 0.0722\,B'$
  - R' – prime denotes gamma correction

$$R' = R^{1/g}$$

- Unitless

Linear vs. gamma-corrected values



Sensitivity to luminance

- Weber-law – the just-noticeable difference is proportional to the magnitude of a stimulus

Ernst Heinrich Weber
[From wikipedia]

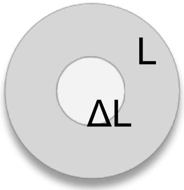
The smallest detectable luminance difference

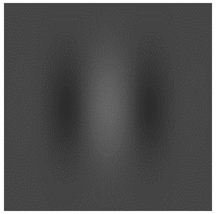
Background (adapting) luminance

$$\frac{\Delta L}{L} = k$$

Constant

Typical stimuli:





Consequence of the Weber-law

- Smallest detectable difference in luminance

$$\frac{\Delta L}{L} = k$$

For k=1%

L	ΔL
100 cd/m^2	1 cd/m^2
1 cd/m^2	0.01 cd/m^2

- Adding or subtracting luminance will have different visual impact depending on the background luminance
- Unlike LDR luma values, luminance values are **not** perceptually uniform!

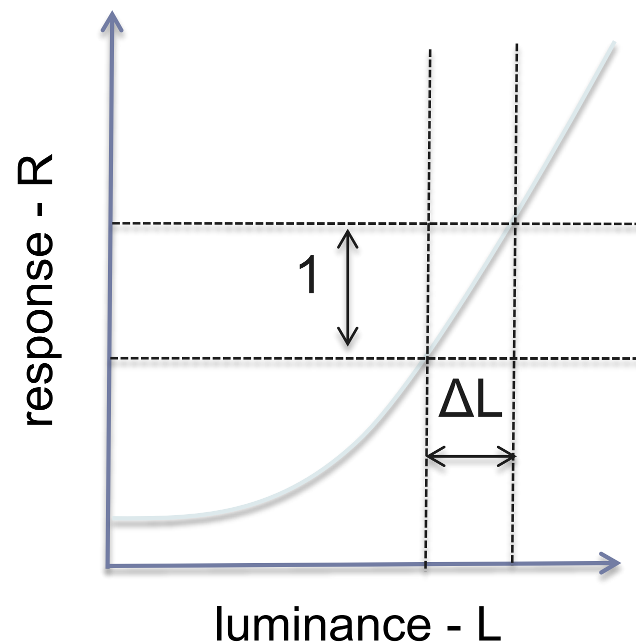
How to make luminance (more) perceptually uniform?

- Using "Fechnerian" integration

$$\frac{dR}{dl}(L) = \frac{1}{\Delta L(L)}$$

Derivative of response

Detection threshold



Luminance transducer:

$$R(L) = \int_0^L \frac{1}{\Delta L(l)} dl$$

Assuming the Weber law

$$\frac{\Delta L}{L} = k$$

- and given the luminance transducer

$$R(L) = \int_0^L \frac{1}{\Delta L(l)} dl$$

- the response of the visual system to light is:

$$R(L) = \int \frac{1}{kL} dL = \frac{1}{k} \ln(L) + k_1$$

Fechner law

$$R(L) = a\ln(L)$$

- Response of the visual system to luminance is **approximately** logarithmic

- The values of HDR pixel values are much more intuitive when they are plotted / considered / processed in the logarithmic domain

Gustav Fechner
[From Wikipedia]

ToC

VR display technologies

TFT-LCD *TN, STN, MVA, PVA, IPS*

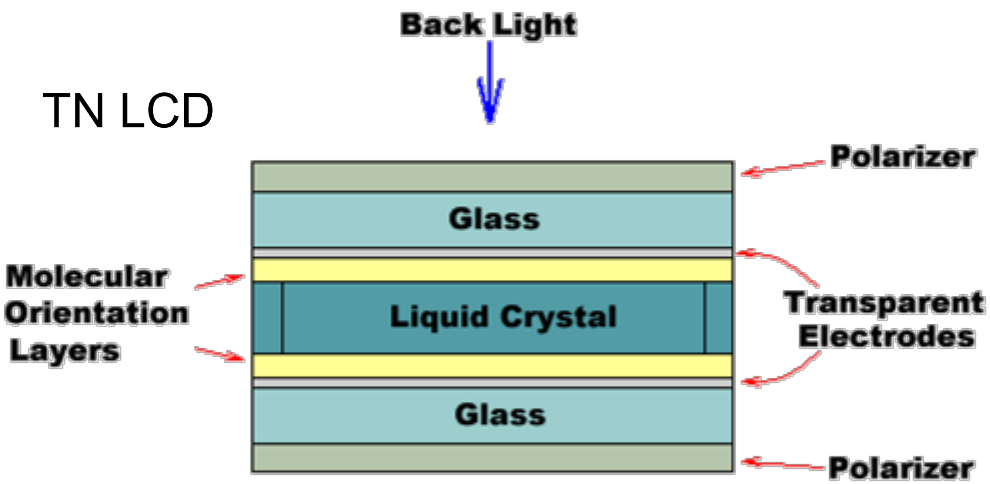
- Contrast: <3000:1
- Transmissive
- Complex temporal response
- Arbitrary bright
- Constant power at constant backlight

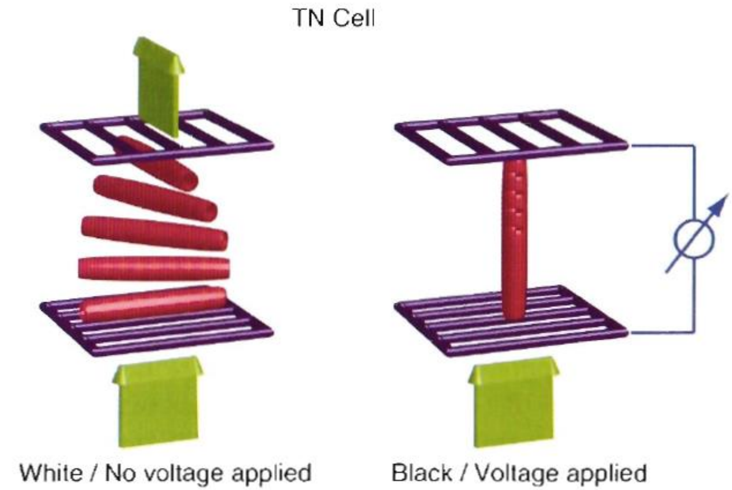
AMOLED

- Contrast: >10,000:1
- Emmisive
- Rapid response
- Brightness affects longevity
- Power varies with image content

LCD

TN LCD

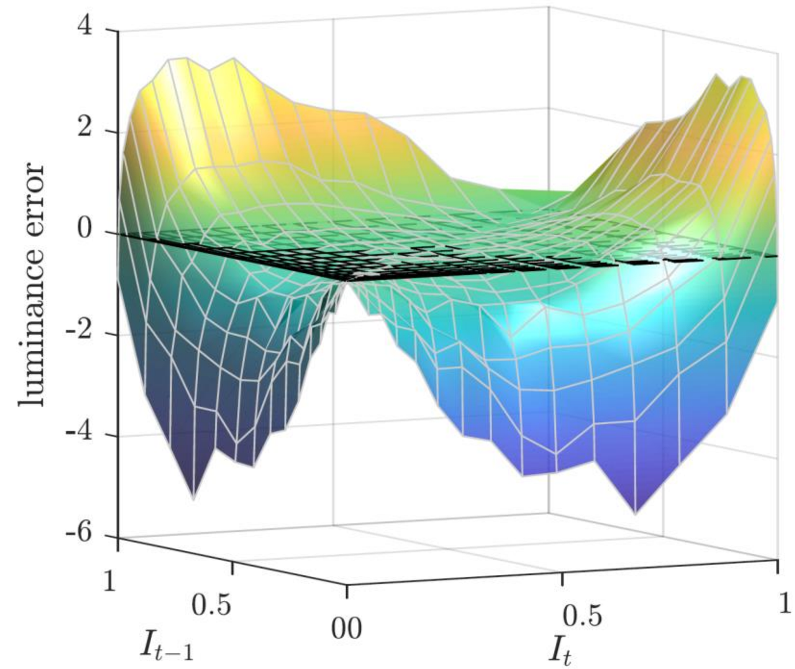


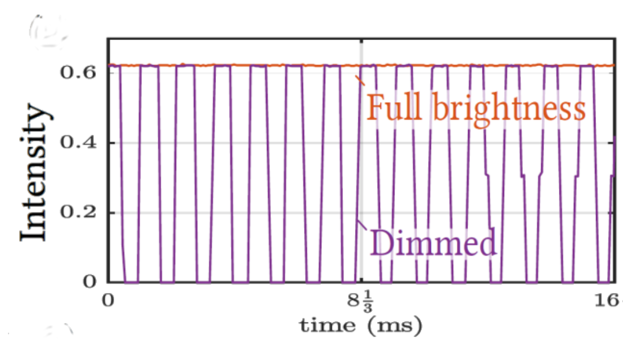


- color may change with the viewing angle
- contrast up to 3000:1
- higher resolution results in smaller fill-factor
- color LCD transmits only up to 8% (more often close to 3-5%) light when set to full white

LCD temporal response

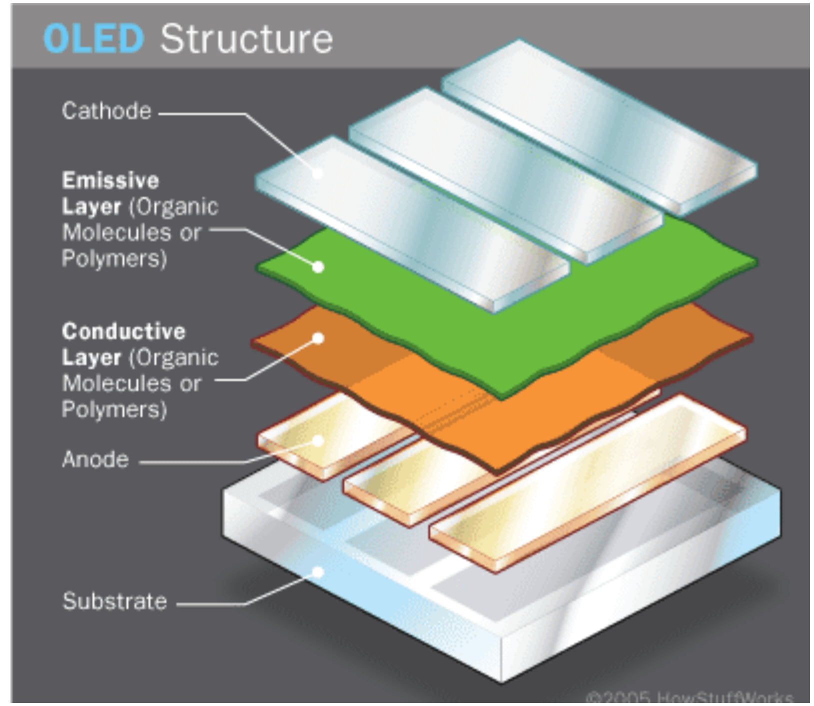
- Experiment on an IPS LCD screen
- We rapidly switched between two intensity levels at 120Hz
- Measured luminance integrated over 1s
- The top plot shows the difference between expected ($\frac{I_{t-1}+I_t}{2}$) and measured luminance
- The bottom plot: intensity measurement for the full brightness and half-brightness display settings





OLED

- based on electrophosphorescence
- large viewing angle
- the power consumption varies with the brightness of the image
- fast (< 1 microsec)
- arbitrary sizes

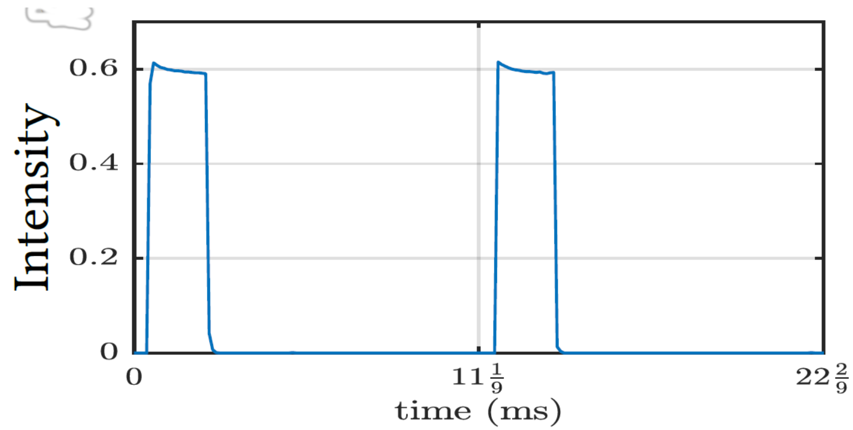


- life-span is a concern
- more difficult to produce

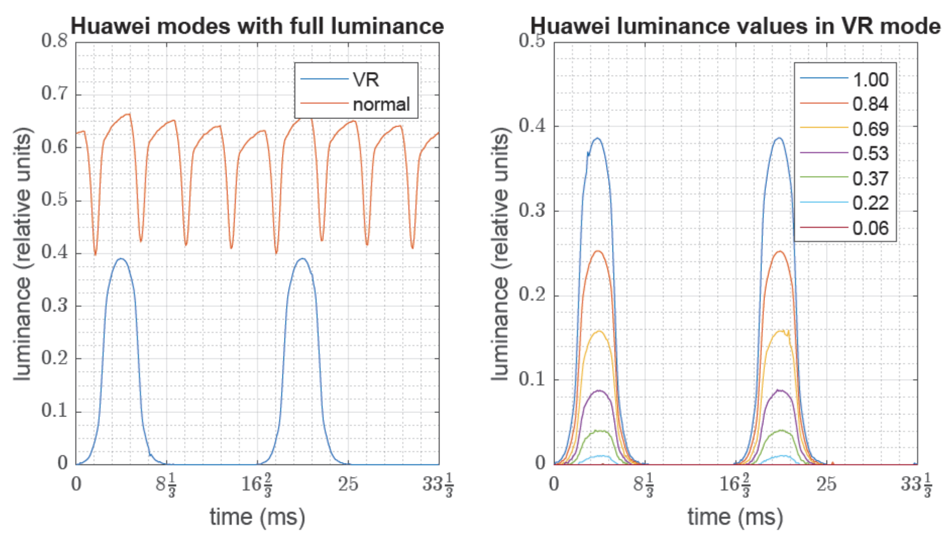
Low persistence displays

- Most VR displays flash an image for a fraction of frame duration
- This reduces hold-type blur
- And also reduces the perceived lag of the rendering

HTC Vive

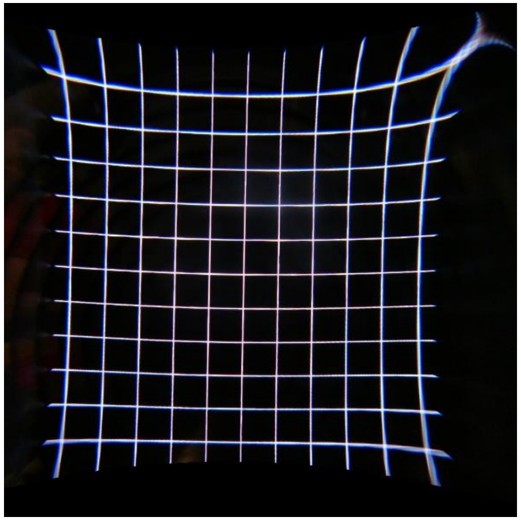


Mate 9 Pro + DayDream



Lens in VR displays

- Aberrations when viewing off-center
  - Chromatic aberration
  - Loss of resolution
  - Difficult to eliminate if the exact eye position is unknown
- Glare
  - Scattering of the light in the lens
  - From Fresnel fringes
  - Reduces dynamic range



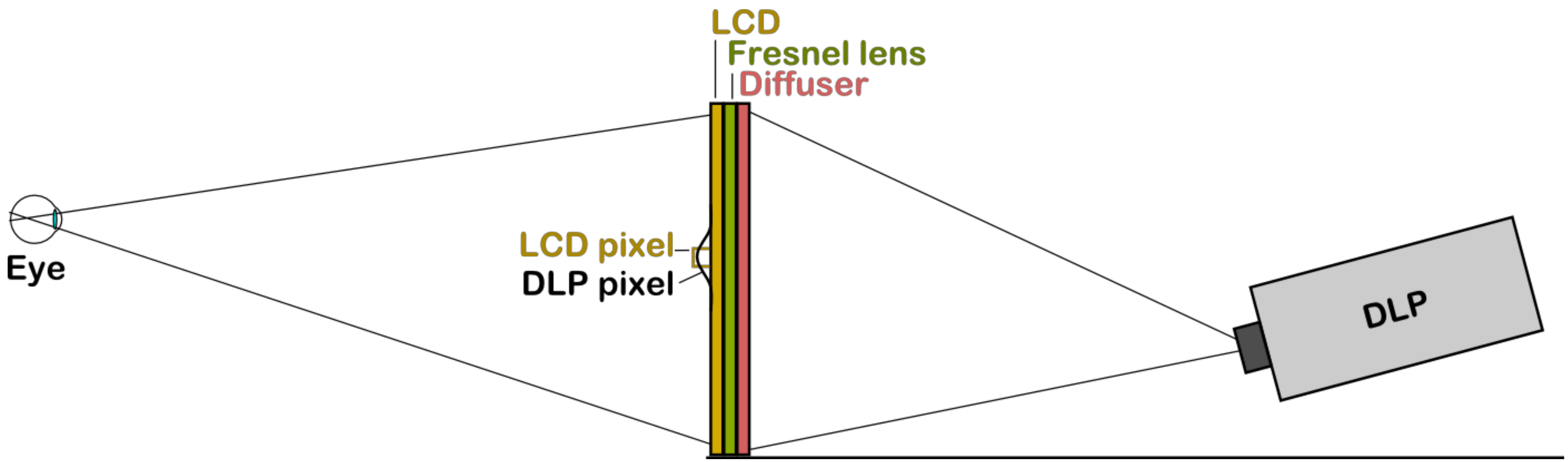
Examples from: http://doc-ok.org/?p=1414

HDR Display

- Modulated LED array
- Conventional LCD
- Image compensation

Low resolution LED Array x High resolution Colour Image = High Dynamic Range Display

HDR display



Desired image

DLP blur (PSF)

$$\arg\min_{L,D} ||I - g\, D \circ L||_2$$

DLP image

LCD image

Subject to:

$$\forall(x,y) \quad L_{min} \leq L(x,y) \leq L_{max}$$

$$\forall(x,y) \quad D_{min} \leq D(x,y) \leq D_{max}$$

Resolution

- Relevant units: pixels per visual degree [ppd]
- Nyquist frequency in cycles per degree = ½ of ppd
- PC & mobile resolution
  - 1981: 12” 320x200 monitor @50cm: 10.9 ppd
  - 1990: 12” 1024x768 monitor @50cm: 37 ppd
  - 2011: 3.5” 960x640 iPhone @30cm: 68 ppd
  - 2016: 31” 4K monitor @50cm: 50 ppd
  - 2018: 6” phone @30cm: 117 ppd
- VR resolution
  - 2016 HTC Vive: 10 ppd
  - 2018 HTC Vive Pro: 13 ppd

ToC

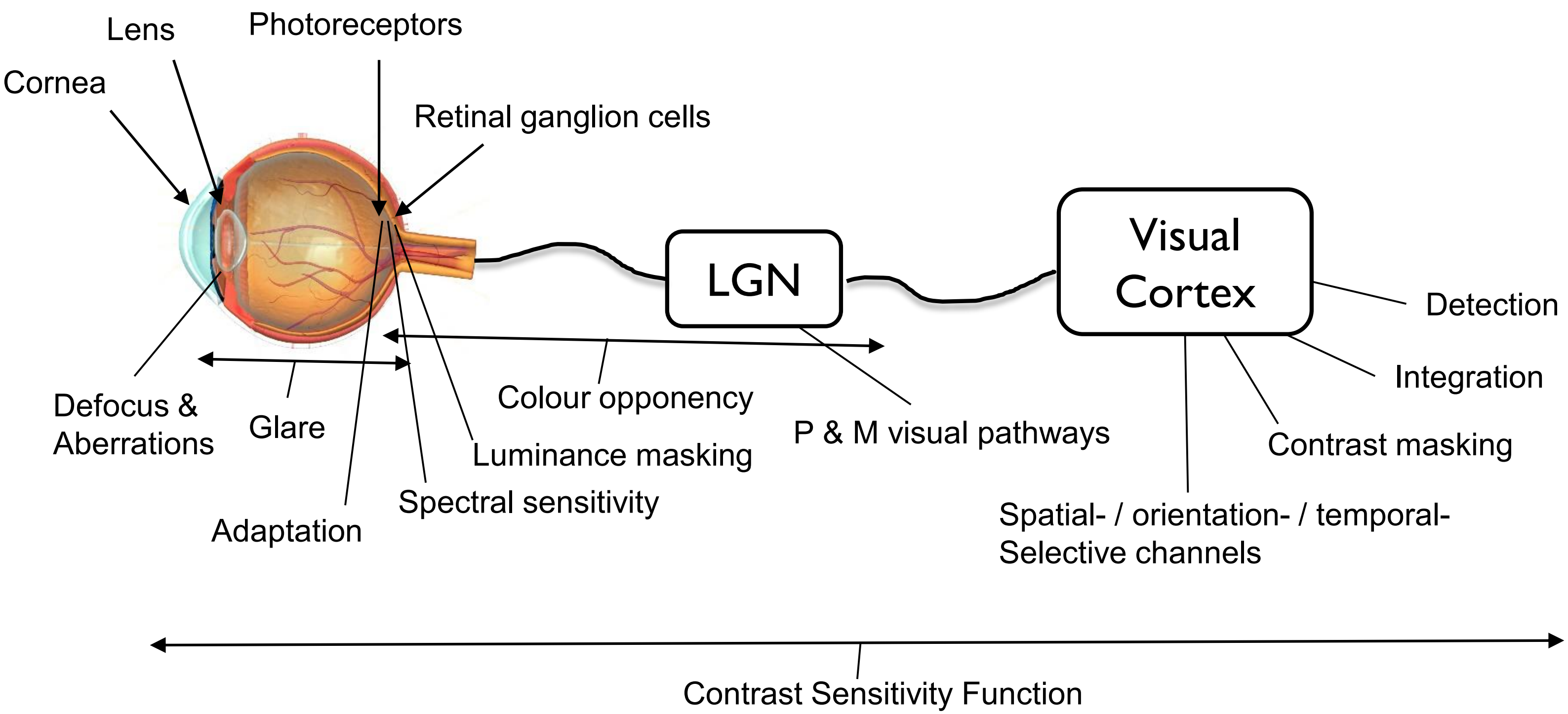
- HDR in a nutshell
- Display technologies in VR
- Perception & image quality
- Example: Temporal Resolution Multiplexing

(Camera) image reconstruction model

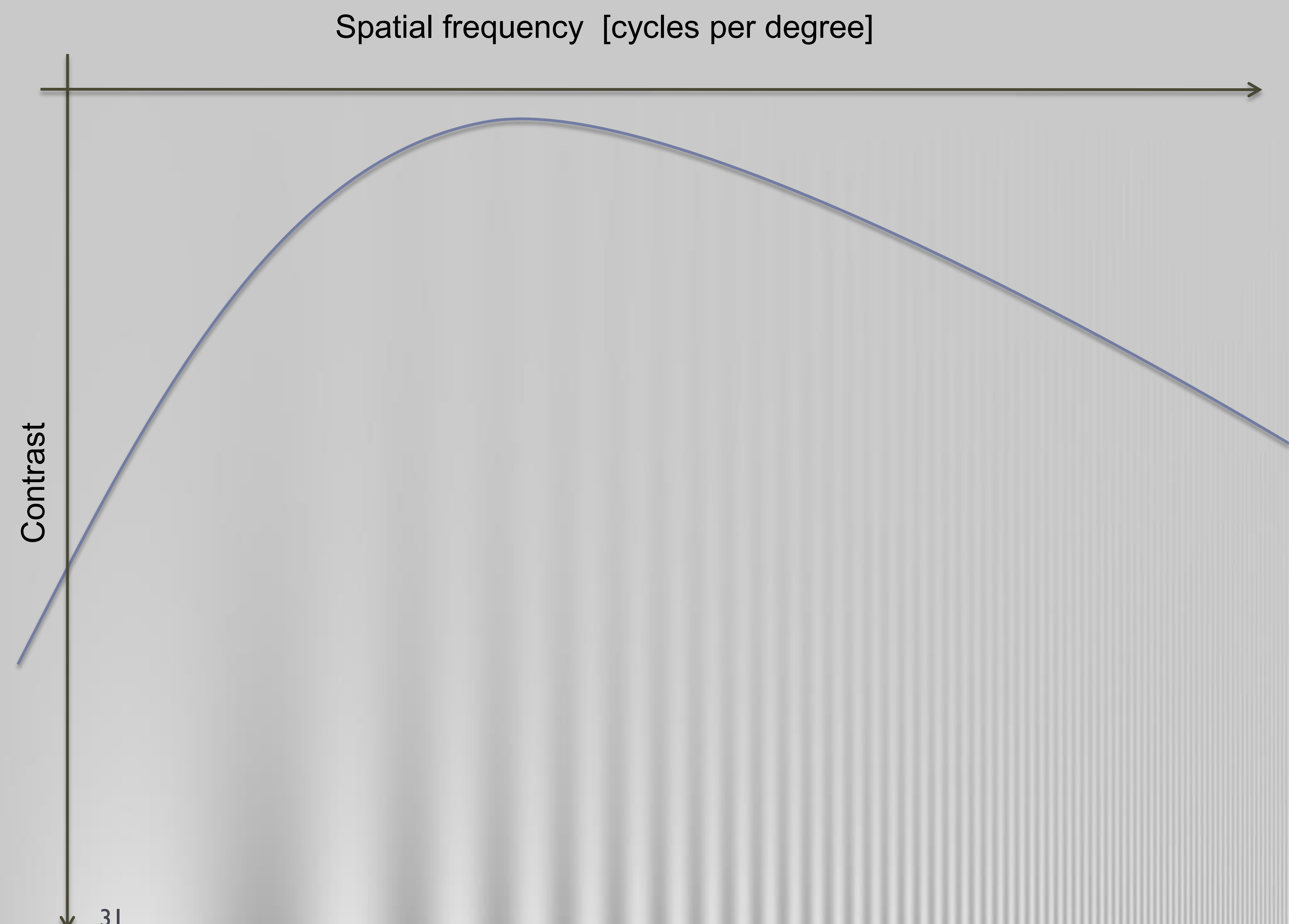
$$\overbrace{Y}^{\text{Captured image}} = \underbrace{g}_{\text{Convolution kernel}}\overbrace{X}^{\text{Latent image}} + \underbrace{\eta}_{\text{Noise}}$$

- Can we come up with a similar model for visual system?

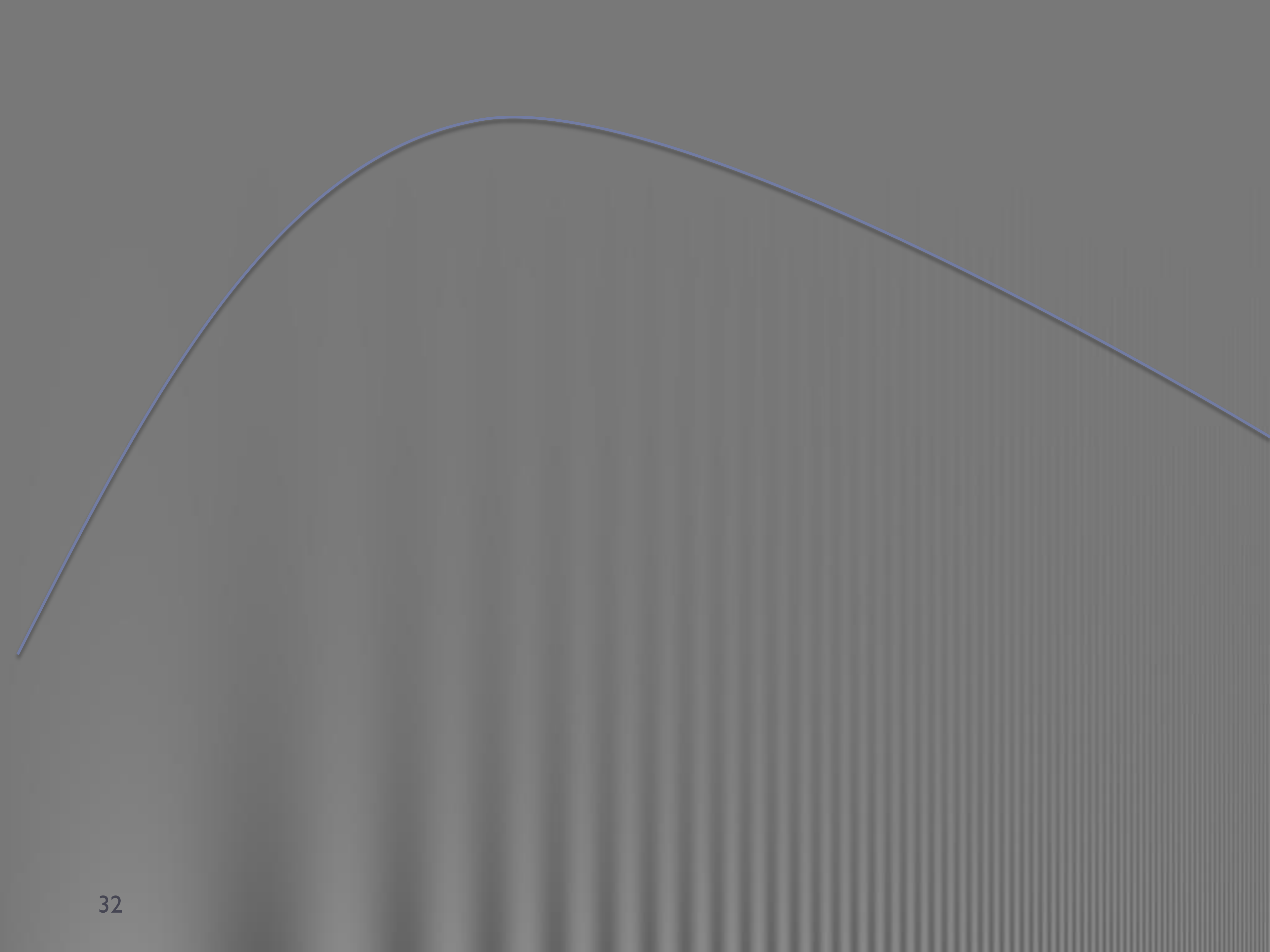
Modeling visual system



Excellent visualization of the human eye:
https://animagraffs.com/human-eye/



Campbell & Robson contrast sensitivity chart



Contrast Sensitivity Function

Temporal frequency

Stimulus size

Spatial frequency

Eccentricity

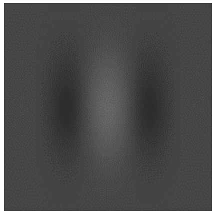
$$CSF = S(r, q, w, l, i^2, d, e)$$

Orientation

Viewing distance

Adapting luminance

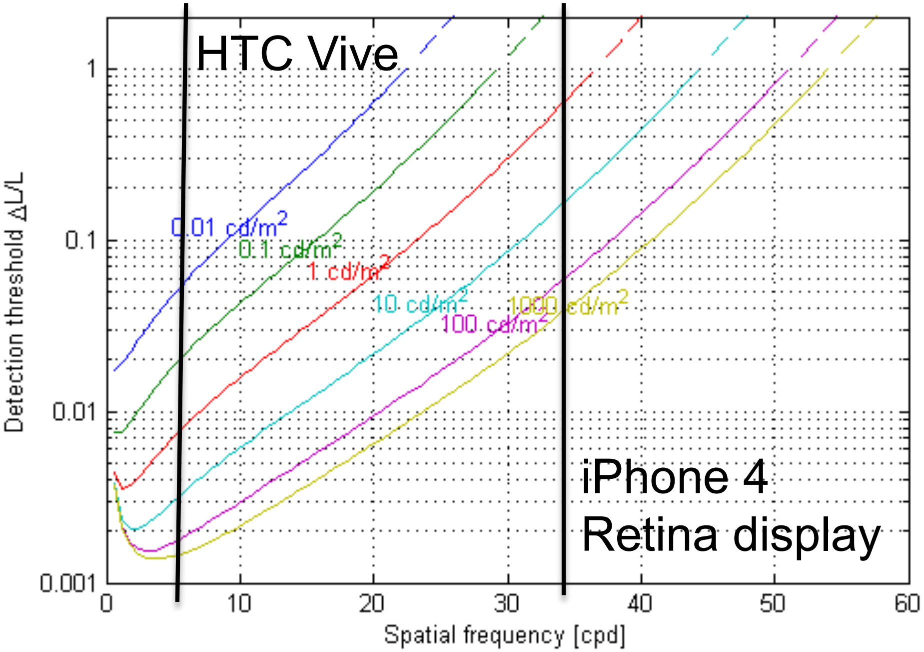
Contrast Sensitivity Function



- Sensitivity = inverse of the detection threshold

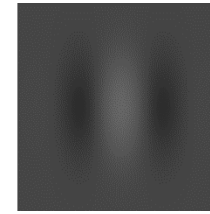
$$S = \frac{L_b}{\Delta L}$$

- Detection of barely noticeable luminance difference ΔL on a uniform background L_b
- Varies with luminance

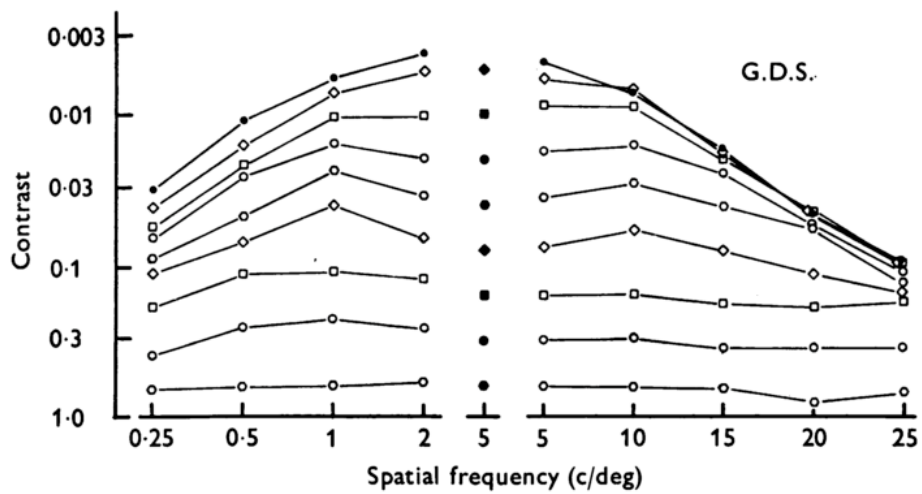


CSF models:
Barten, P. G. J. (2004). https://doi.org/10.1117/12.537476
Mantiuk, R., Kim, K. J., Rempel, A. G., & Heidrich, W. (2011) https://doi.org/10.1145/2010324.1964935

Contrast Constancy



- CSF is NOT MTF of visual system
  - Contrast constancy
  - There is little variation in magnitude of perceived contrast above the detection threshold



Georgeson and Sullivan. 1975. J. Phsysio

Modeling visual perception

- Since visual system is highly non-linear, a linear model

$$Y = gX + \eta$$

CSF is NOT MTF!

cannot be used.

- Visual processing is an unknown non-linear function:

Input image

Percept
(not an image)

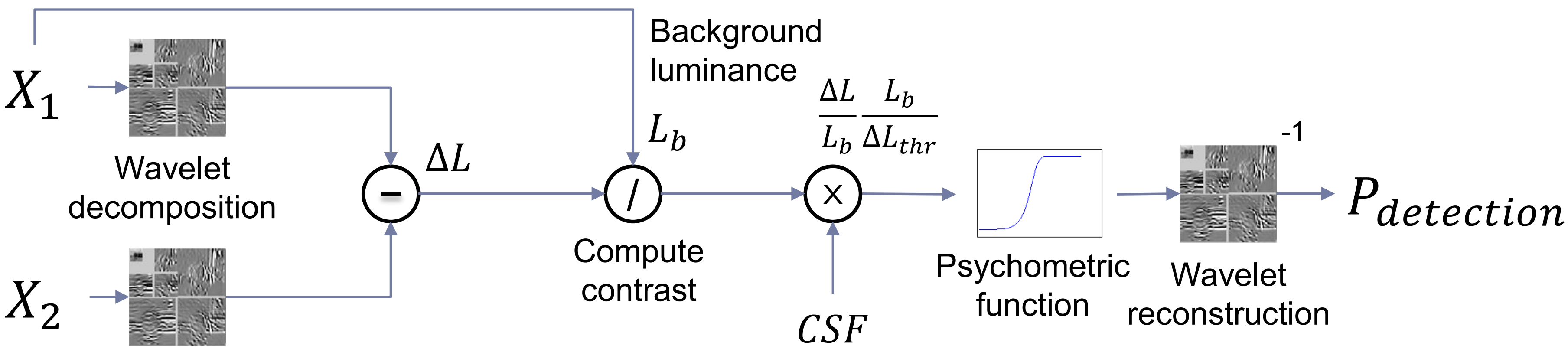
$$Y = f[X]$$

Visual processing

Predicting visible differences with CSF

- But we can use CSF to find the probability of spotting a difference beween a pair of images X_1 and X_2:

$$p(f[X_1] = f[X_2] \,|X_1, X_2, CSF)$$

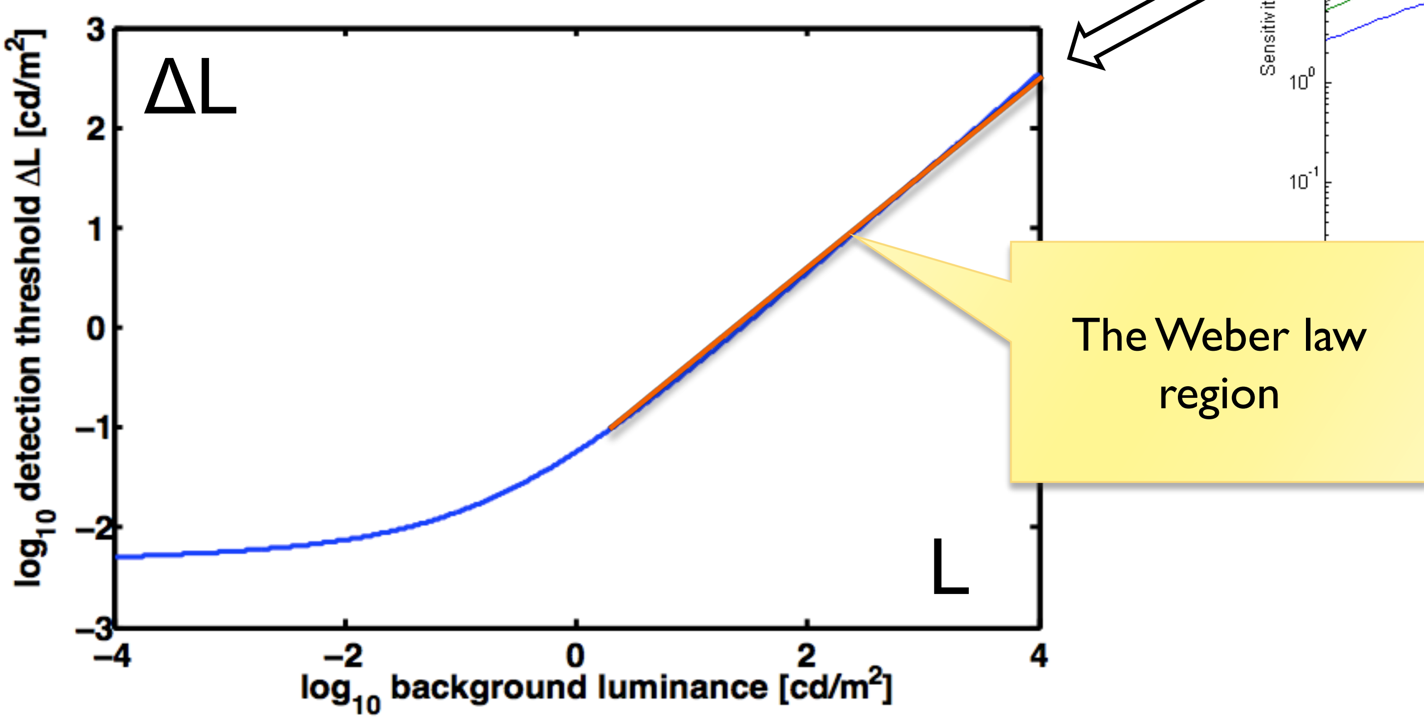


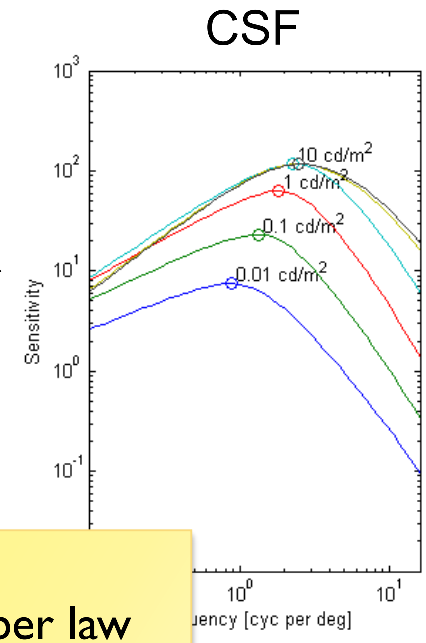
(simplified) Visual Difference Predictor

Daly, S. (1993).
Mantiuk, R., et al. (2011)
https://doi.org/10.1145/2010324.1964935

Fixing Fechner law

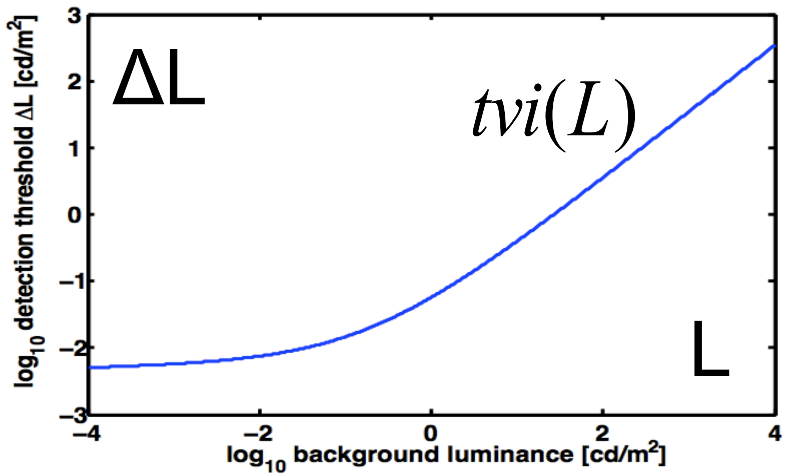
- Peaks of the CSF across luminance
  - The most conservative threshold





Weber-law revisited

- If we allow detection threshold to vary with luminance according to the t.v.i. function:

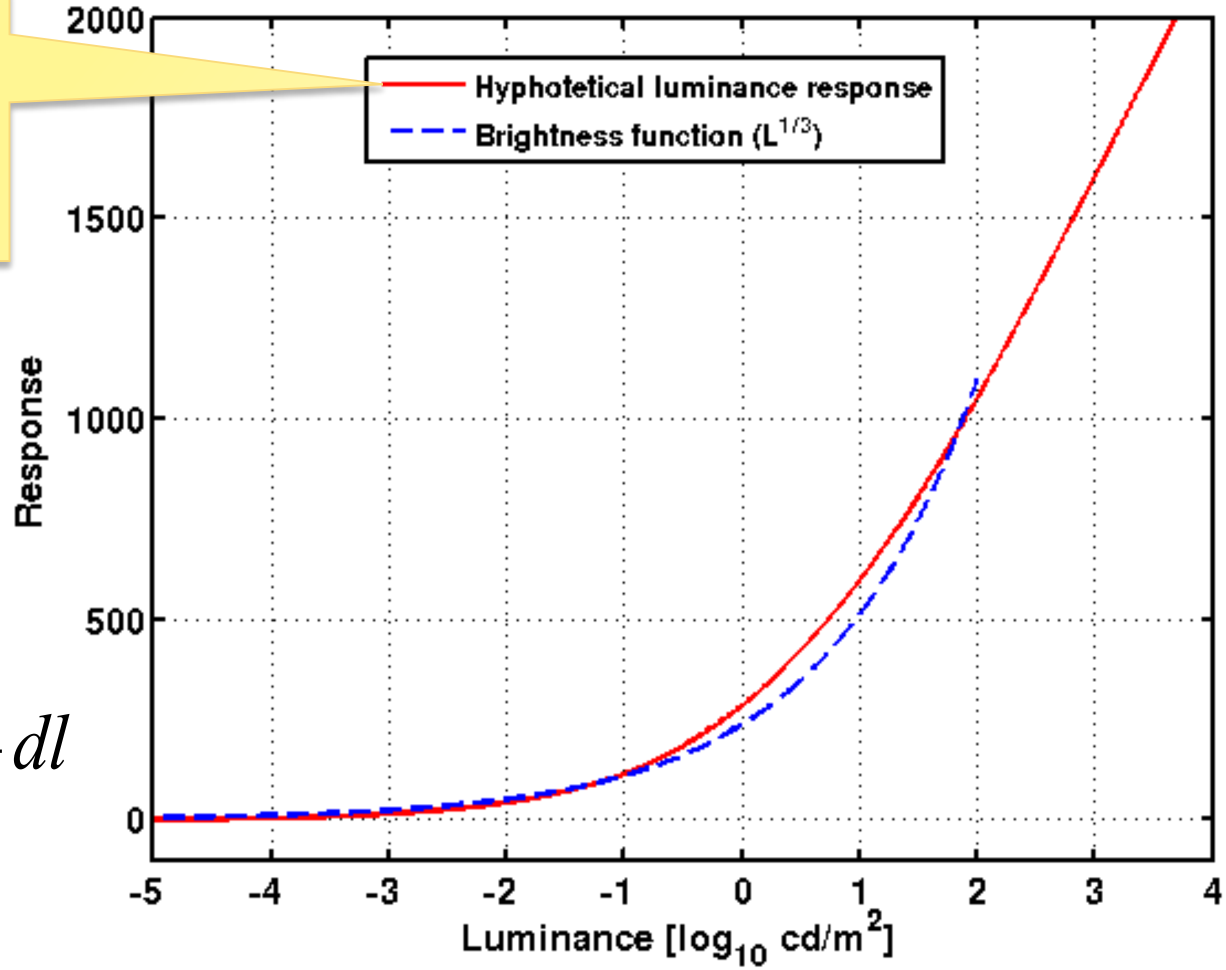


- we can get more accurate estimate of the "response":

$$R(L) = \int_0^L \frac{1}{tvi(l)}\, dl$$

Fechnerian integration and Stevens' law

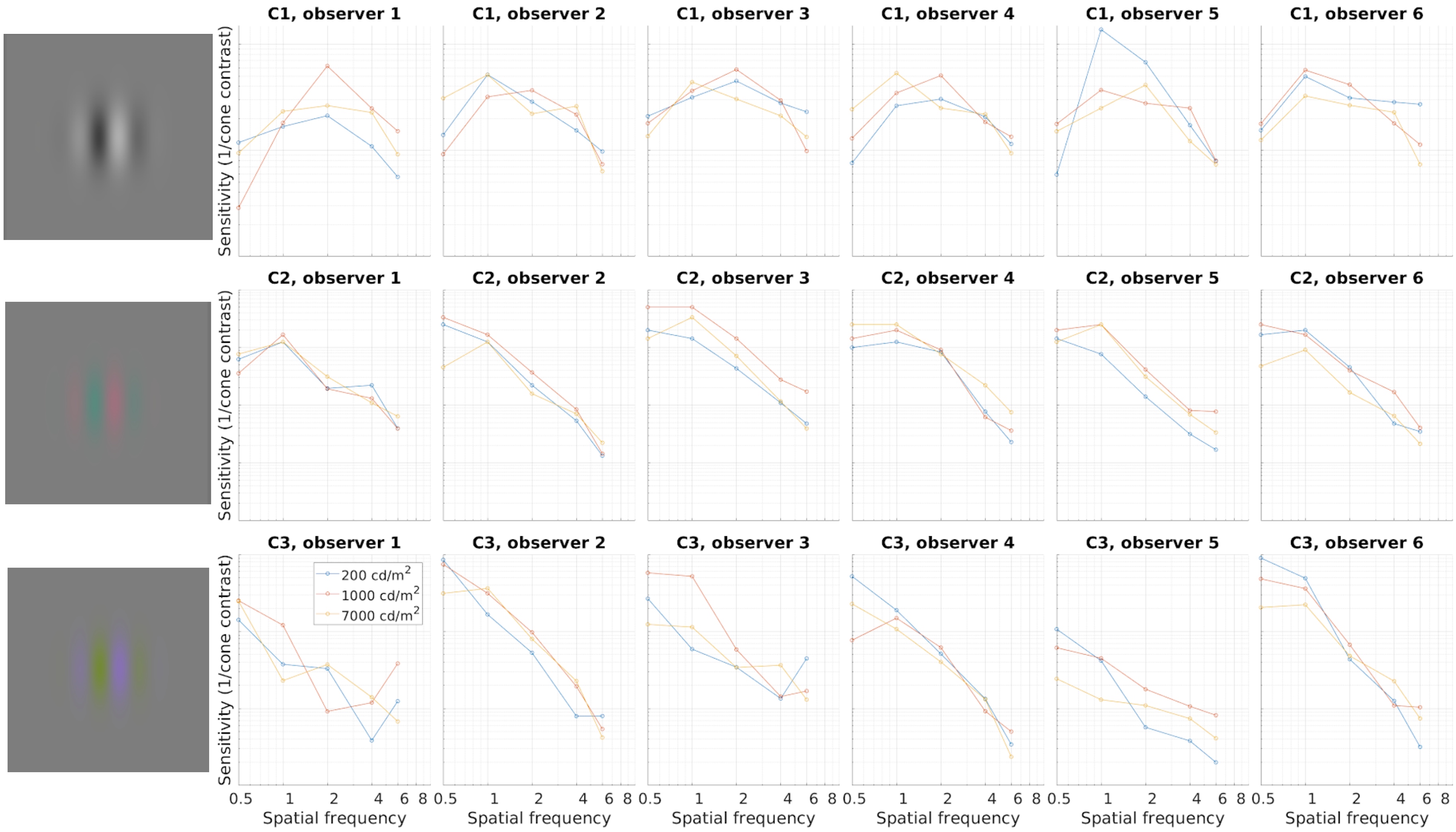
R(L) - function derived from the t.v.i. function



$$R(L) = \int_0^L \frac{1}{tvi(l)}\,dl$$

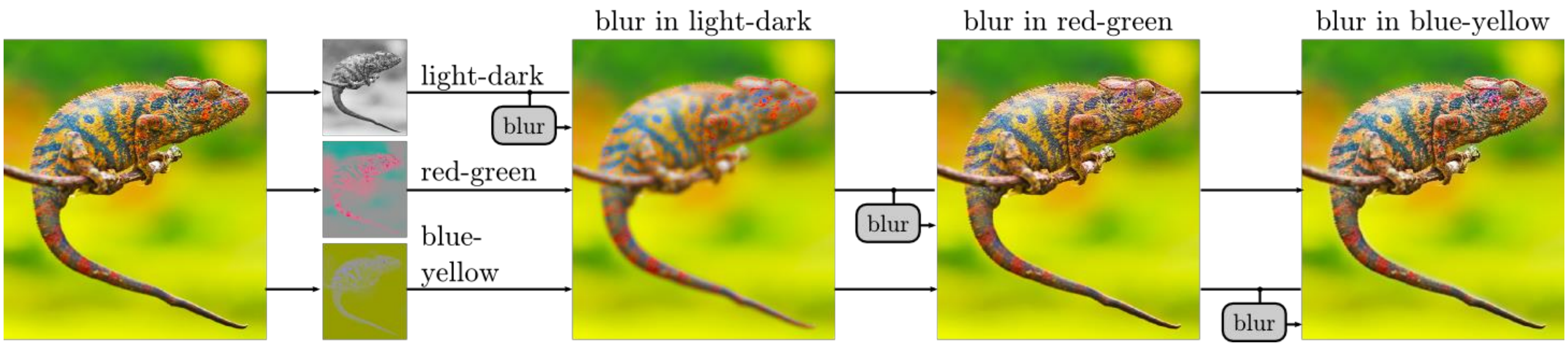
Spatio-chromatic CSF

Per-observer results – fixed cycles



Spatio-chromatic CSF

- Chromatic channels (red-green, blue-yellow) are much less sensitive to high frequencies

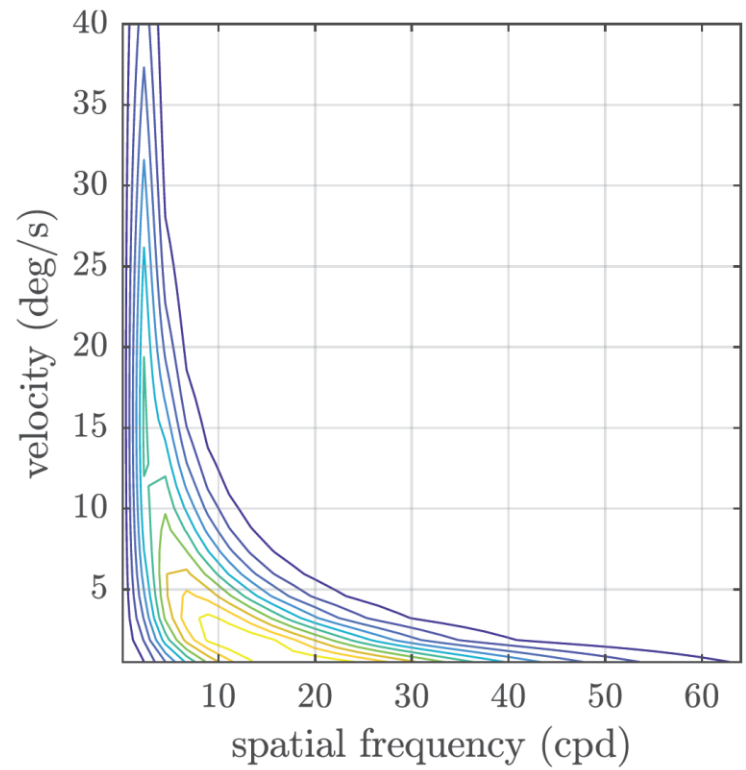


- This is why we can (often) get away with chroma subsampling in image/video compression

Retinal velocity

- Sensitivity drops rapidly once images start to move
- The eye tracks moving objects
  - Smooth Pursuit Eye Motion (SPEM)
  - Stabilizes images on the retina
  - But tracking is not perfect
- Loss of sensitivity mostly caused by imperfect SPEM
  - SPEM worse at high velocities
- Motion sharpenning
  - Relatively small effect

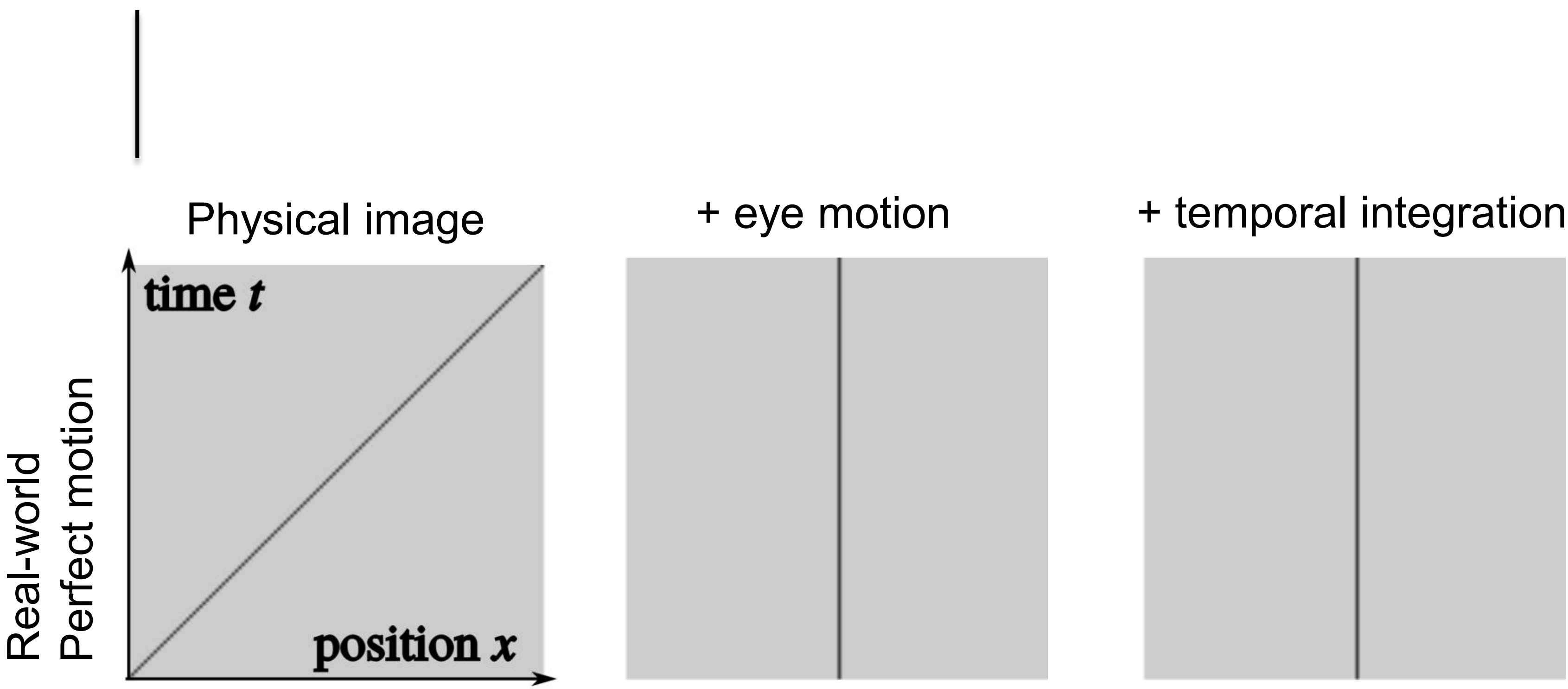
Spatio-velocity contrast sensitivity



Kelly's model [1979]

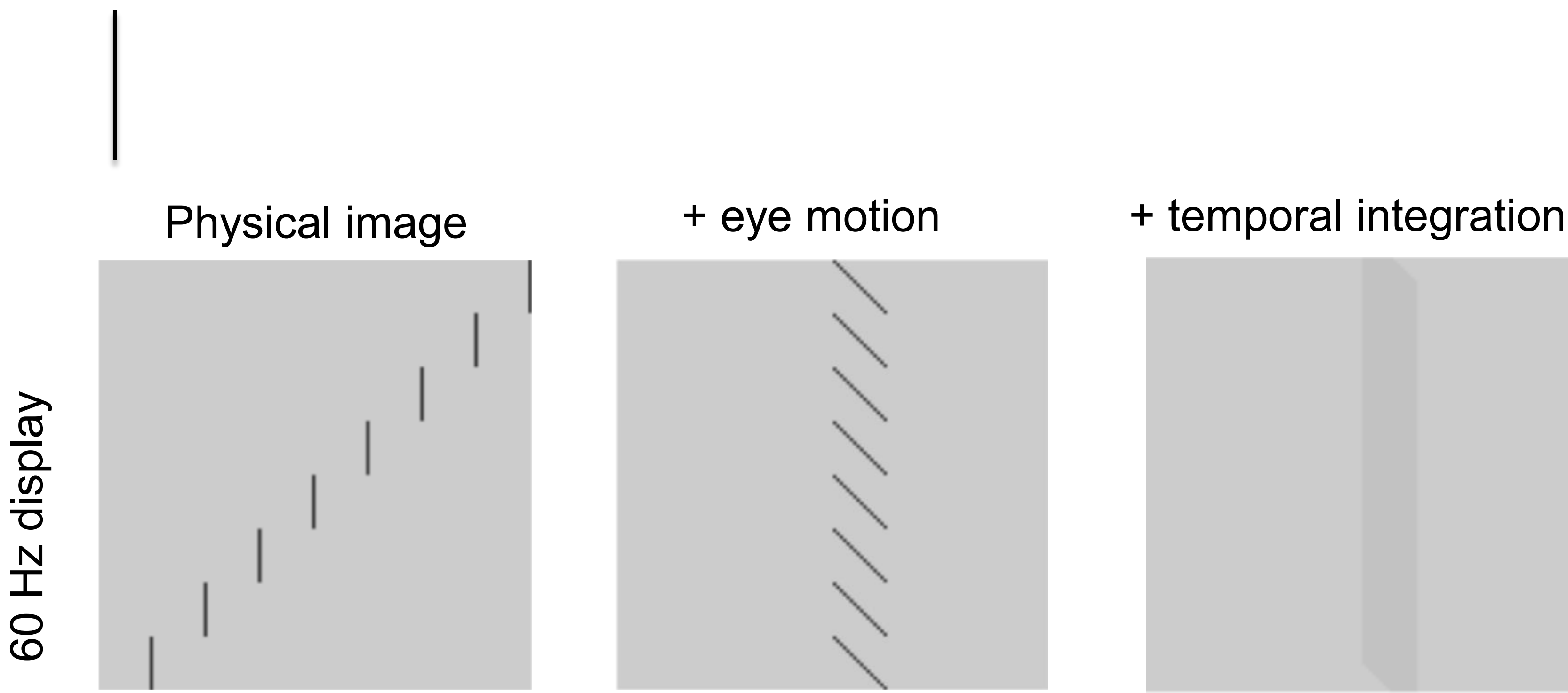
Hold-on blur

- The eye smoothly follows a moving object
- But the image on the display is "frozen" for 1/60th of a second



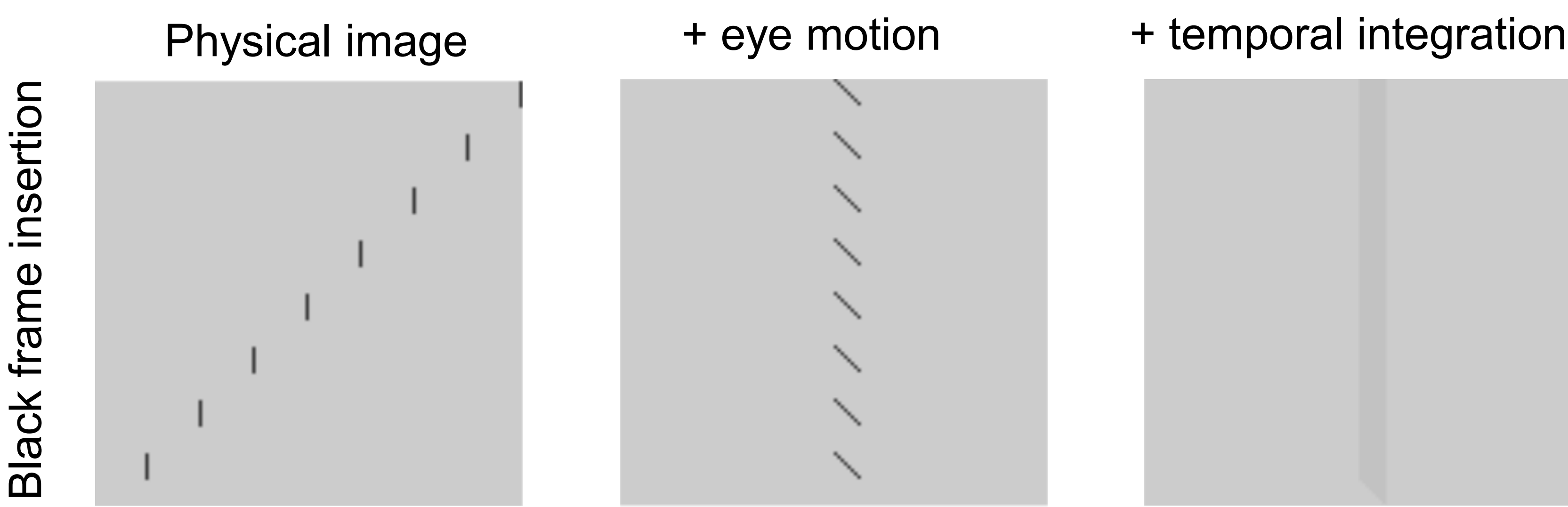
Hold-on blur

- The eye smoothly follows a moving object
- But the image on the display is "frozen" for 1/60th of a second



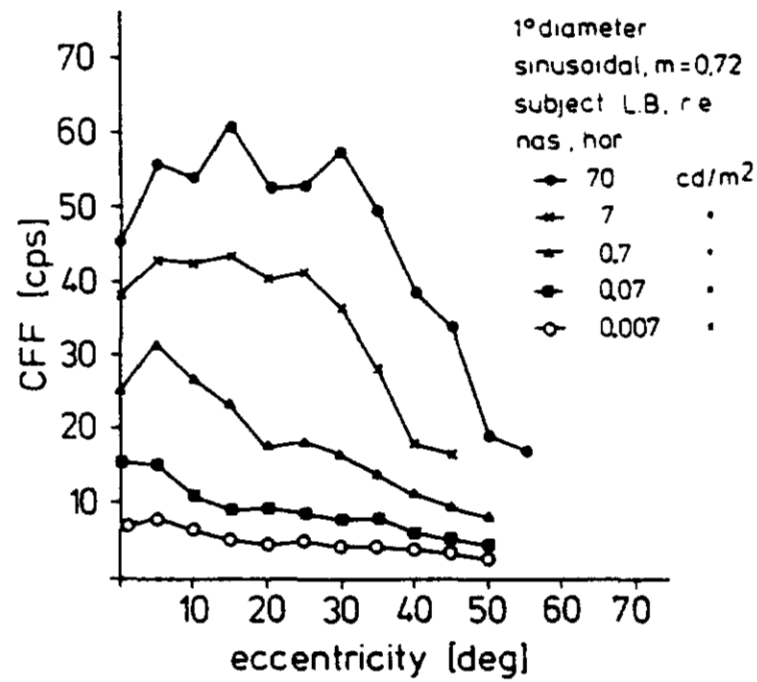
Hold-on blur

- The eye smoothly follows a moving object
- But the image on the display is "frozen" for 1/60th of a second



Flicker

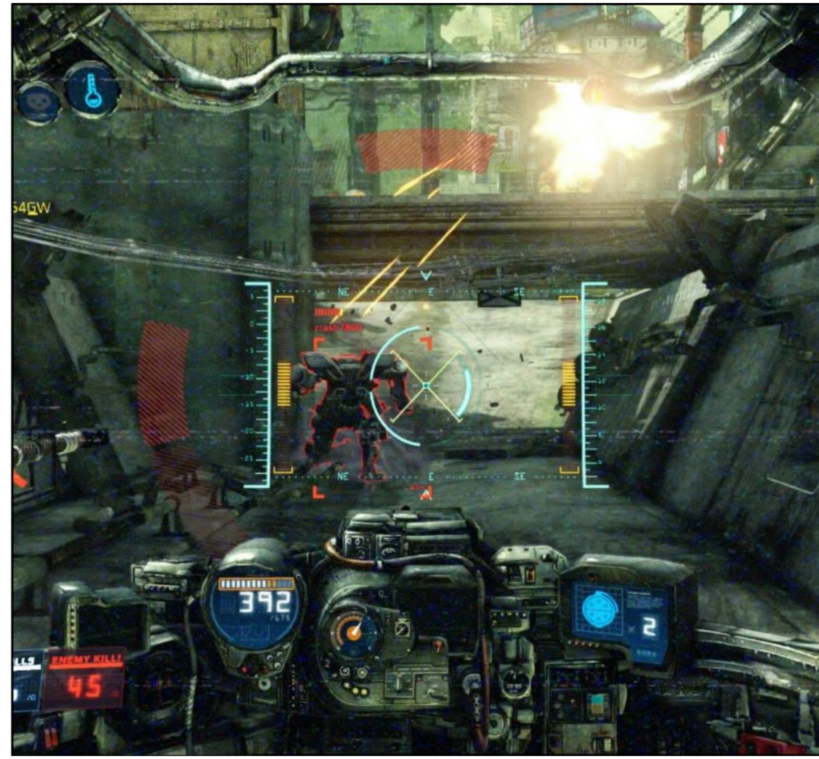
- Critical Flicker Frequency
  - Strongly depends on luminance – big issue for HDR VR headsets
  - Increases with eccentricity
  - and stimulus size
  - It is possible to detect flicker even at 2kHz
    - For saccadic eye motion



[Hartmann et al. 1979]

Simulation sickness

- Conflict between vestibular and visual systems
  - When camera motion inconsistent with head motion
  - Frame of reference (e.g. cockpit) helps
  - Worse with larger FOV
  - Worse with high luminance and flicker



ToC

- HDR in a nutshell
- Display technologies in VR
- Perception & image quality
- Example: Temporal Resolution Multiplexing

VR rendering – required bandwidth

$$\underset{\text{2 eyes}}{\underset{\uparrow}{2}} \times \underbrace{(1400 \times 1600)}_{\text{resolution}} \times \underset{\text{refresh rate}}{\underset{\uparrow}{90}} \times \underset{\text{pixel data}}{\underset{\uparrow}{3}} \approx 1.13GB \approx 9\text{Gbps}$$

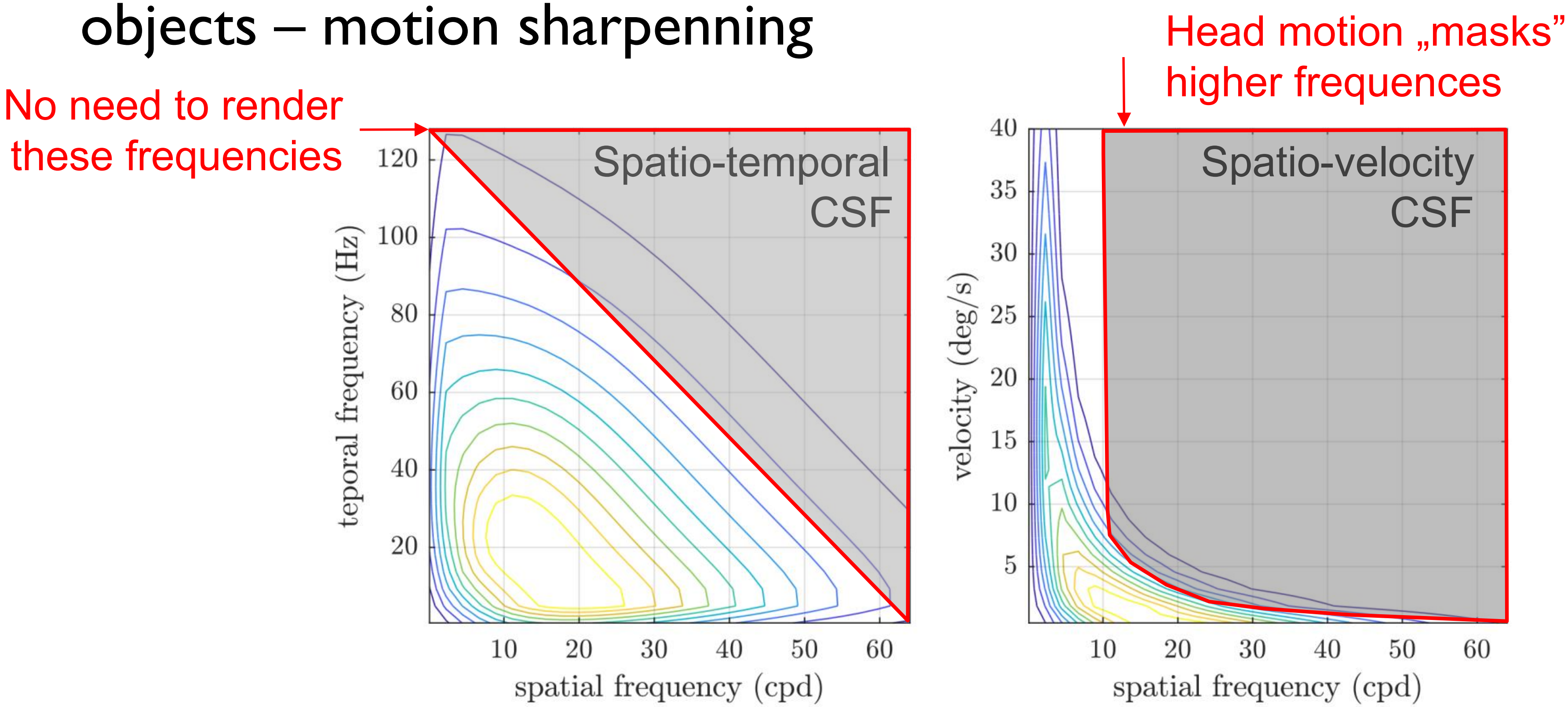
TRM: Temporal Resolution Multiplexing

- Render every second frame at a lower resolution
- Transfer high- and low-resolution frames
- When displaying
  - Compensate for the loss of high frequencies
  - Model display and its limitations
  - Handle the limited dynamic range

See the demo in the break!

TRM: Why does it work?

- The eye cannot see high spatio-temporal frequencies
- The eye cannot see the loss of sharpness for moving objects – motion sharpenning



Summary

- VR/AR display technologies must exploit the limitations of the visual system
  - Because the display / rendering bandwidth is becoming too large
- HDR for VR is a great idea because
  - It gives more realistic experience
  - Better quality with the same number of pixels
  - Additional depth cues
- HDR for VR is bad idea because
  - Increased flicker visibility
  - Increased simulation sickness
  - Lens glare will reduce effective dynamic range

References

- Concise overview of high dynamic range imaging
  - Mantiuk, R. K., Myszkowski, K., & Seidel, H. (2015). High Dynamic Range Imaging. In *Wiley Encyclopedia of Electrical and Electronics Engineering* (pp. 1–42). Hoboken, NJ, USA: John Wiley & Sons, Inc. https://doi.org/10.1002/047134608X.W8265
  - Downloadable PDF: http://www.cl.cam.ac.uk/~rkm38/pdfs/mantiuk15hdri.pdf
- Comprehensive book on display technologies
  - Hainich, R. R., & Bimber, O. (2011). *Displays: Fundamentals and Applications*. CRC Press.
  - https://goo.gl/RLe8nA
- Book on HDR Imaging
  - Reinhard, E., Heidrich, W., Debevec, P., Pattanaik, S., Ward, G., & Myszkowski, K. (2010). High Dynamic Range Imaging: Acquisition, Display, and Image-Based Lighting (2nd editio). Morgan Kaufmann.
- Computational models of visual perception
  - WANDELL, B.A. 1995. *Foundations of vision*. Sinauer Associates.

SIGGRAPH ASIA 2018 TOKYO

Christian Richardt

Motion-Aware Displays

SIGGRAPH Asia Course on Cutting-Edge VR/AR Display Technologies

CAMERA
Centre for the Analysis of Motion, Entertainment Research and Applications

UNIVERSITY OF BATH

richardt.name
c_richardt

Why care about motion?

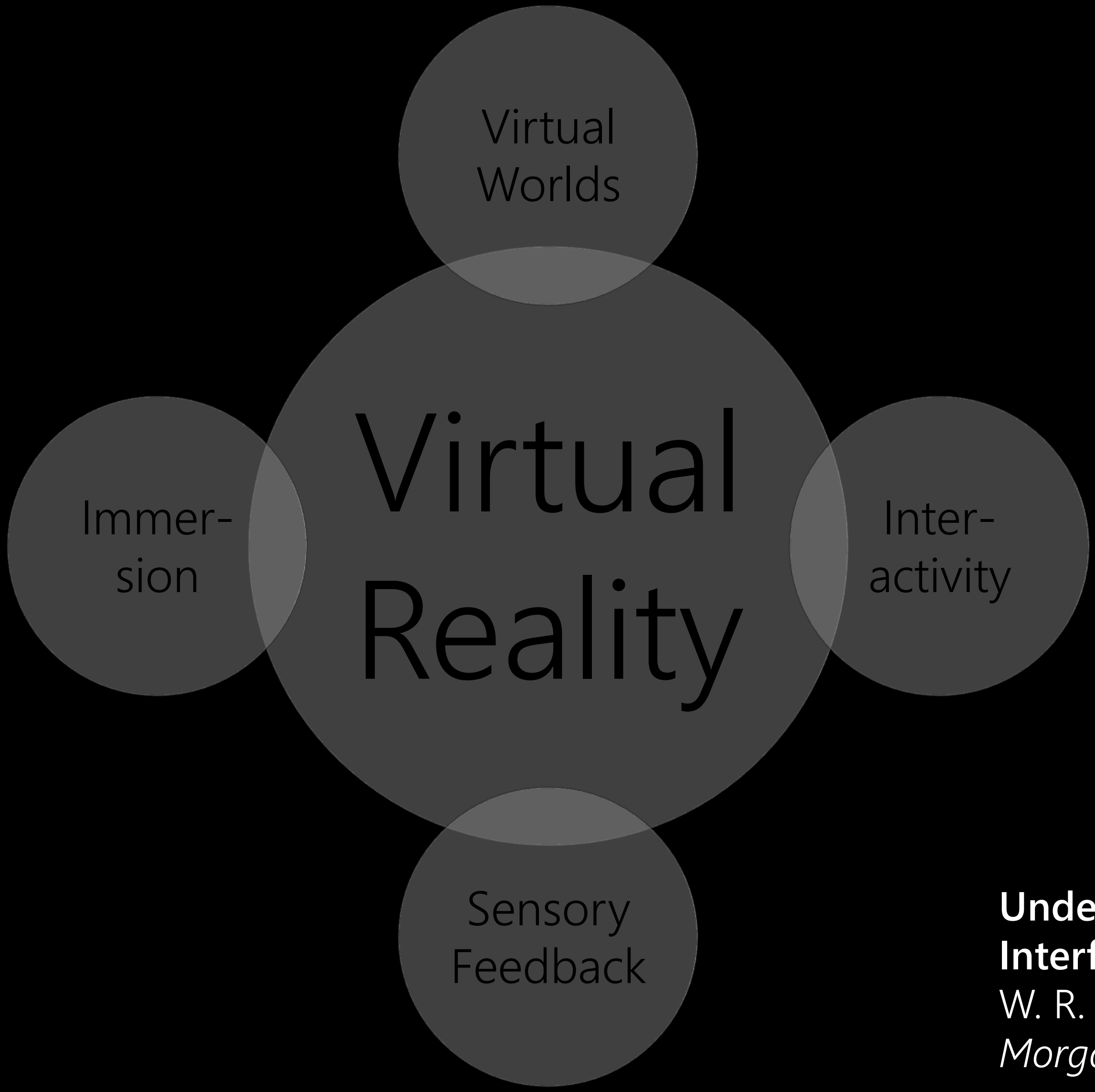
The world's first VR HMD by Ivan Sutherland (1968): Miniature CRTs, head tracking with mechanical sensors (in the video, "Sword of Damocles") or ultrasonic sensors

- Need to track motion to generate the right images:
  - head motion
  - hand motion
  - full-body motion
- Motion tracking enables:
  - **immersion** = the replacement of perception with virtual stimuli
  - **presence** = the sensation of "being there"

Motion-aware displays

1. Perception of immersion
2. Tracking in VR and AR
3. Hand input devices
4. Motion capture

Virtual reality experiences



Understanding Virtual Reality: Interface, Application, and Design
W. R. Sherman & A. B. Craig
Morgan Kaufmann Publishers, 2003

Immersion vs Presence

- **Immersion** is an objective notion which can be defined as the sensory stimuli coming from a device, for example a data glove
- Measurable and comparable between devices

- **Presence** is a subjective phenomenon, personal experiences in an immersive environment
- Subjective feeling of being there

A note on presence terminology
M. Slater
Presence Connect, 2003, 3:3

Slide adapted from Zerrin Yumak

Immersion

- sensation of being in another environment
- **Mental immersion:**
  - a movie, game or a novel might immerse you too
  - suspension of disbelief, state of being deeply engaged
- **Physical immersion:**
  - bodily entering into a medium
  - synthetic stimulus of the body’s senses via the use of technology

Slide adapted from Zerrin Yumak

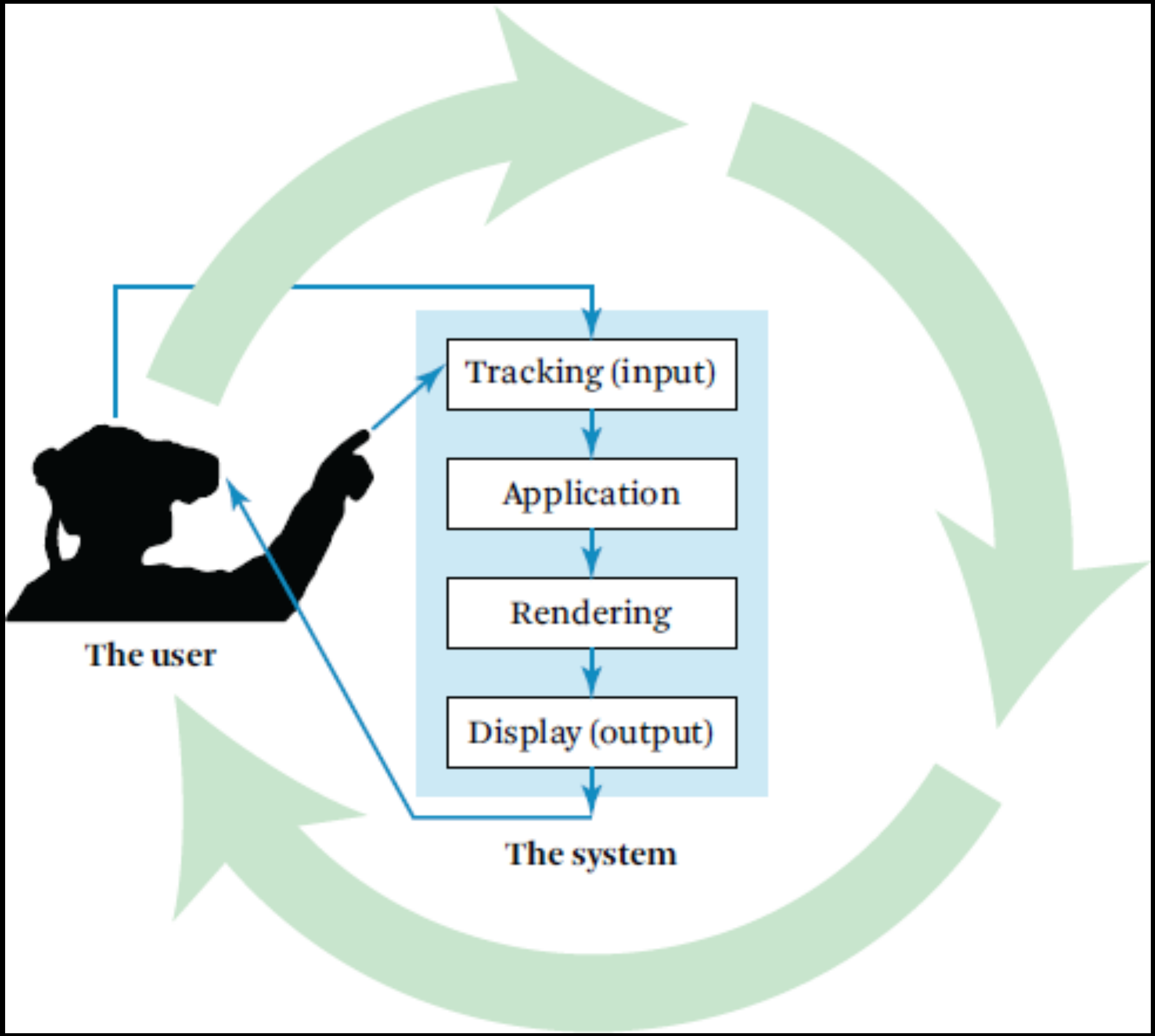
Self-embodiment

- Perception that the user has a body within the virtual world
- The presence of a virtual body can be quite compelling
  - even when that body does not look like one's own body
  - effective for teaching empathy by "walking in someone else's shoes" and can reduce racial bias
- Whereas body shape and colour are not so important, <u>motion is extremely important</u>
- Presence can be broken when visual body motion does not match physical motion

Putting Yourself in the Skin of a Black Avatar Reduces Implicit Racial Bias
T. C. Peck, S. Seinfeld, S. M. Aglioti & M. Slater
Consciousness and Cognition, 2013, 22(3), 779–787

Slide adapted from Zerrin Yumak

VR system input–output cycle

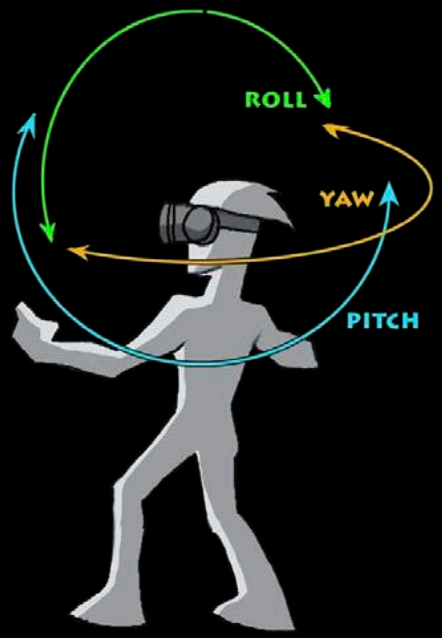


Scene-Motion- and Latency-Perception Thresholds for Head-Mounted Displays
J. J. Jerald
PhD Thesis, UNC Chapel Hill, 2009

Tracking degrees of freedom (DoF)

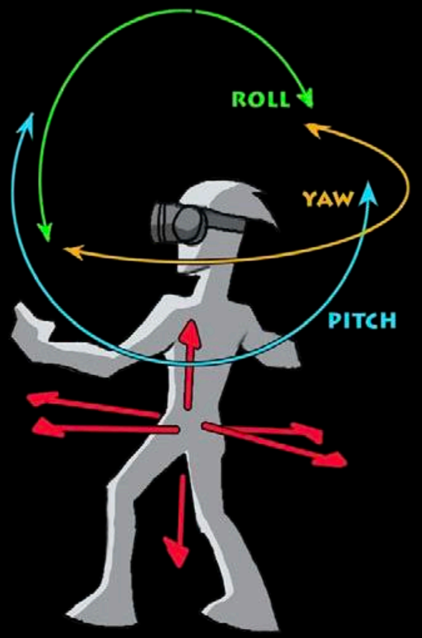
3 degrees of freedom (3-DoF)

- "In which direction am I looking"
- Detect rotational head movement
- Look around the virtual world from a fixed point



6 degrees of freedom (6-DoF)

- "Where am I and in which direction am I looking"
- Detect rotations and translational movement
- Move in the virtual world like in the real world



Slide adapted from Qualcomm Technologies, Inc.

Tracking technologies

- Mechanical:
  - e.g. physical linkage
- Electromagnetic:
  - e.g. magnetic sensing
- Inertial:
  - e.g. accelerometers, MEMs
- Acoustic:
  - e.g. ultrasonic
- Optical:
  - computer vision
- Hybrid:
  - combination of technologies

contact-less tracking

Slide adapted from Bruce Thomas & Mark Billinghurst

Mechanical tracking

- Idea: mechanical arms with joint sensors
- Advantages:
  - high accuracy
  - low jitter
  - low latency
- Disadvantages:
  - cumbersome
  - limited range
  - fixed position

Ivan Sutherland's Sword of Damocles (1968)

MicroScribe (2005)

Slide adapted from Bruce Thomas & Mark Billinghurst

Magnetic tracking

- Idea: measure difference in current between a magnetic transmitter and a receiver
- Advantages:
  - 6-DoF, robust & accurate
  - no line of sight needed
- Disadvantages:
  - limited range, noisy
  - sensitive to metal
  - expensive

Razer Hydra (2011)
Magnetic source with two wired controllers
short range (<1 m), precision of 1 mm and 1°
62 Hz sampling rate, <50 ms latency

Magic Leap One (2018)
Transmitter generates 3 orthogonal magnetic fields; unknown specs

Slide adapted from Bruce Thomas & Mark Billinghurst

Inertial tracking

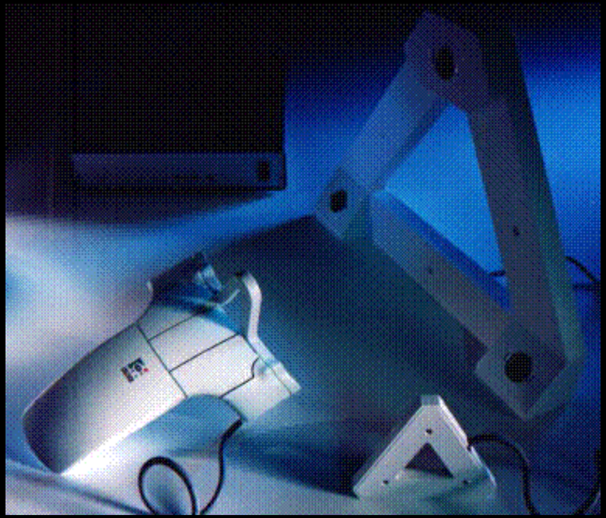
- Idea: Measuring linear and angular orientation rates (accelerometer/gyroscope)
- Advantages:
  - no transmitter, wireless
  - cheap + small
  - high sample rate
- Disadvantages:
  - drift + noise
  - only 3-DoF

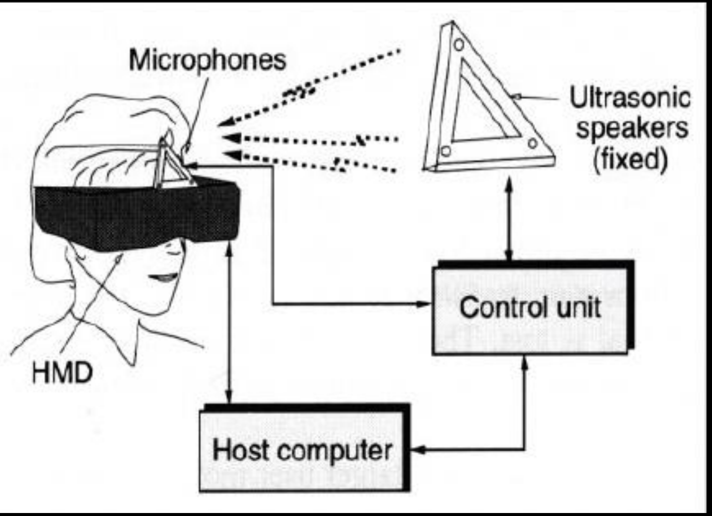
Google Daydream View (2017)
relies on the phone for processing and tracking
3-DoF rotational only tracking of phone + controller

Slide adapted from Bruce Thomas & Mark Billinghurst

Acoustic tracking

- Idea: time-of-flight or phase-coherent sound waves
- Advantages:
  – small + cheap
- Disadvantages:
  – only 3-DoF
  – low resolution
  – low sampling rate
  – requires line-of-sight
  – affected by environment (pressure, temperature)





Logitech 3D Head Tracker (1992)
Transmitter has 3 ultrasonic speakers, 30 cm apart; receiver has 3 mics
range: ~1.5 m, accuracy: 0.1° orientation, 2% distance
50 Hz update, 30 ms latency

Slide adapted from Bruce Thomas & Mark Billinghurst

Optical tracking

- Idea: image processing and computer vision to the rescue
- often using infrared light, retro-reflective markers, multiple views
- Advantages:
  - long range, cheap
  - immune to metal
  - usually very accurate
- Disadvantages:
  - requires markers, line of sight
  - can have low sampling rate

Microsoft Kinect (2010)
IR laser speckle projector, RGB + IR cameras
range: 1–6 m, accuracy: <5 mm
30 Hz update rate, 100 ms latency

Slide adapted from Bruce Thomas & Mark Billinghurst

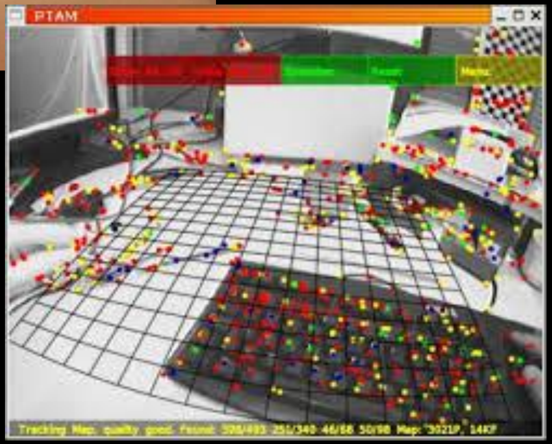
AR optical tracking

- Marker tracking:
  - tracking known artificial images
    - e.g. ARToolKit square markers
- Markerless tracking:
  - tracking from known features in real world
    - e.g. Vuforia image tracking
- Unprepared tracking:
  - in unknown environments
    - e.g. SLAM (simultaneous localisation and mapping)

devfun-lab.com

mobilegeeks.de

PTAM



Slide adapted from Bruce Thomas & Mark Billinghurst

Hybrid tracking

- Idea: multiple technologies overcome limitations of each one
- A system that utilizes two or more position/orientation measurement technologies (e.g. inertial + visual)
- Advantages:
  - robust
  - reduce latency
  - increase accuracy
- Disadvantages:
  - more complex + expensive

digitaltrends.com

Apple ARKit (2017), Google ARCore (2018)
visual-inertial odometry – combine inertial motion sensing with feature point tracking

Slide adapted from Bruce Thomas & Mark Billinghurst

Example: Vive Lighthouse tracking

- Outside-in hybrid tracking:
  - 2 base stations: each with 2 laser scanners, LED array
- Headworn/handheld sensors:
  - 37 photo sensors in HMD, 17 in hand
  - additional IMU sensors (500 Hz)
- Performance:
  - tracking fuses sensor samples at 250 Hz
  - 2 mm RMS accuracy
  - large area: 5×5 m² range
- See: https://youtu.be/xrsUMEbLtOs

gizmodo.com

slashgear.com

Slide adapted from Bruce Thomas & Mark Billinghurst

Hand input devices

- Devices that integrate hand input into VR:
  - world-grounded input devices
  - non-tracked handheld controllers
  - tracked handheld controllers
  - hand-worn devices
  - hand tracking

digitaltrends.com

Slide adapted from Bruce Thomas & Mark Billinghurst

World-grounded hand input devices

- Devices constrained or fixed in the real world
  - e.g. joysticks, steering wheels
- Not ideal for VR
  - constrains user motion
- Good for VR vehicle metaphor, location-based entertainment
  - e.g. driving simulators, Disney's "Aladdin's Magic Carpet Ride"

aliexpress.com

realityprime.com

Slide adapted from Bruce Thomas & Mark Billinghurst

Non-tracked handheld controllers

- Devices held in hand
  - buttons
  - joysticks
  - game controllers
- Traditional video game controllers
  - e.g. Xbox controller

Bottomless Joystick
katsumotoy.com/bj/

techadvisor.co.uk

Slide adapted from Bruce Thomas & Mark Billinghurst

Tracked handheld controllers

- Handheld controller with 6-DoF tracking
  – combines button/joystick/ trackpad input plus tracking
- One of the best options for VR applications
  – physical prop enhancing VR presence
  – providing proprioceptive, passive haptic touch cues
  – direct mapping to real hand motion

HTC Vive controller

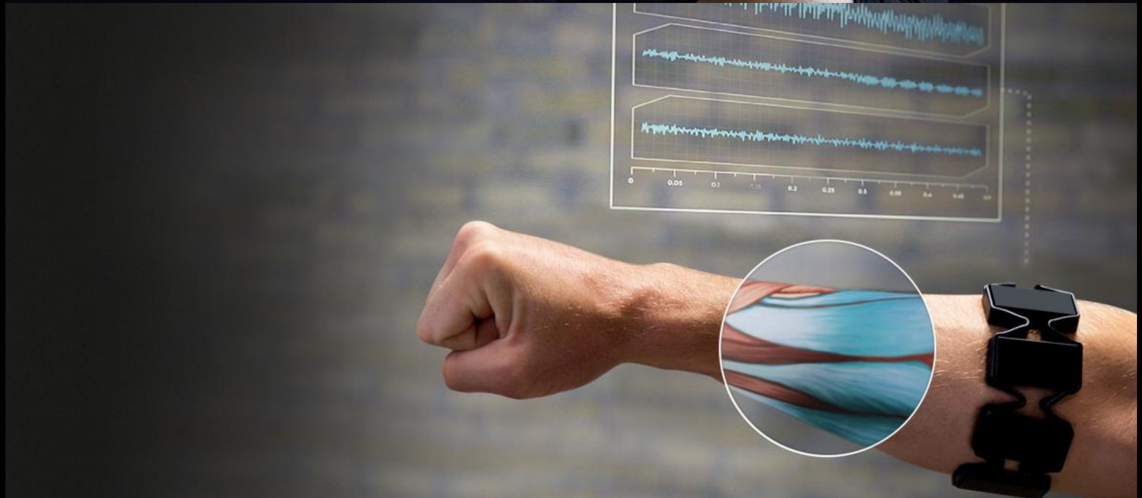
Oculus Touch

Slide adapted from Bruce Thomas & Mark Billinghurst

Hand-worn devices

- Devices worn on hands/arms
  – e.g. glove, EMG sensors, rings
- Advantages:
  – natural input with potentially rich gesture interaction
  – hands can be held in comfortable positions
    - no line-of-sight issues
  – hands and fingers can fully interact with real objects

virtualrealitytimes.com

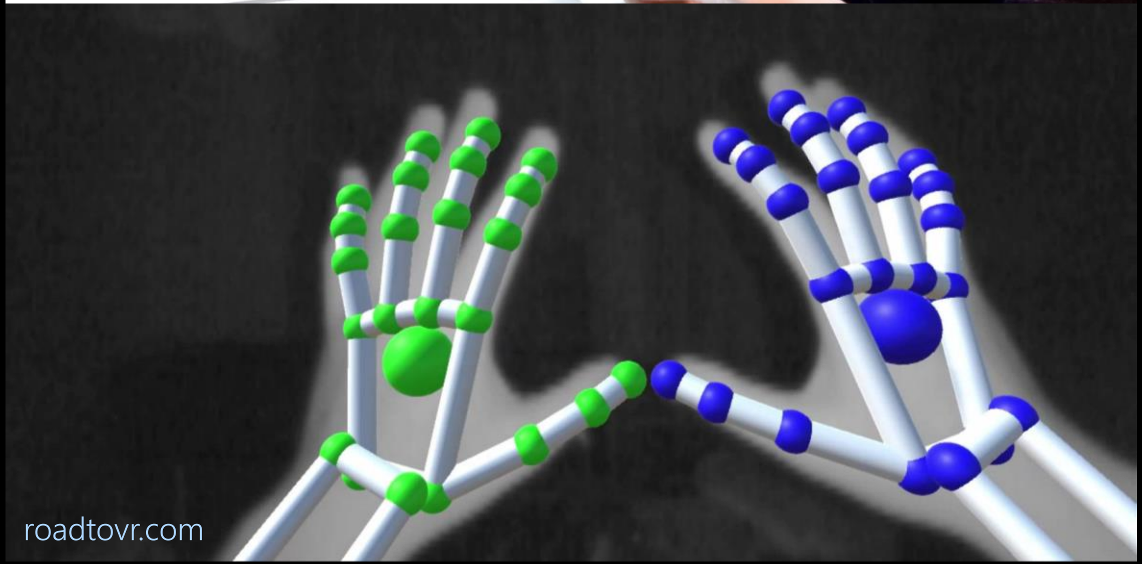


developerblog.myo.com/raw-uncut-drops-today/

Slide adapted from Bruce Thomas & Mark Billinghurst

Hand tracking

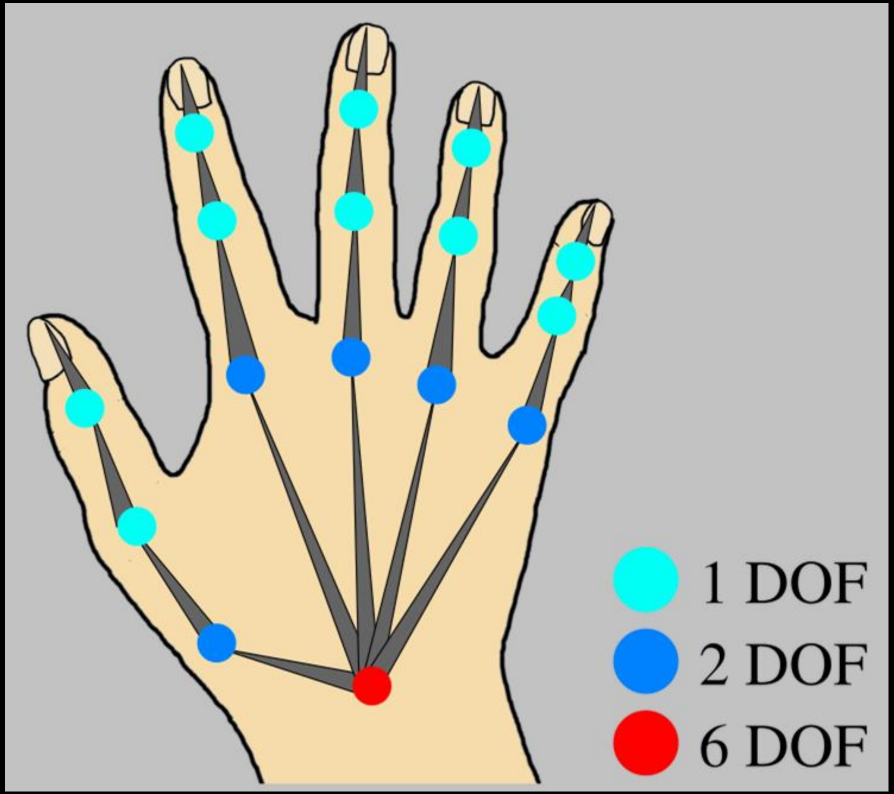
- Using computer vision to track bare hand input
- Creates compelling sense of presence, natural interaction
- Advantages:
  - least intrusive, purely passive
  - hands-free tracking, so can interact freely with real objects
  - low power requirements, cheap
  - more ubiquitous, works outdoors



Slide adapted from Bruce Thomas & Mark Billinghurst, Fransizka Mueller

Case study: Egocentric hand tracking

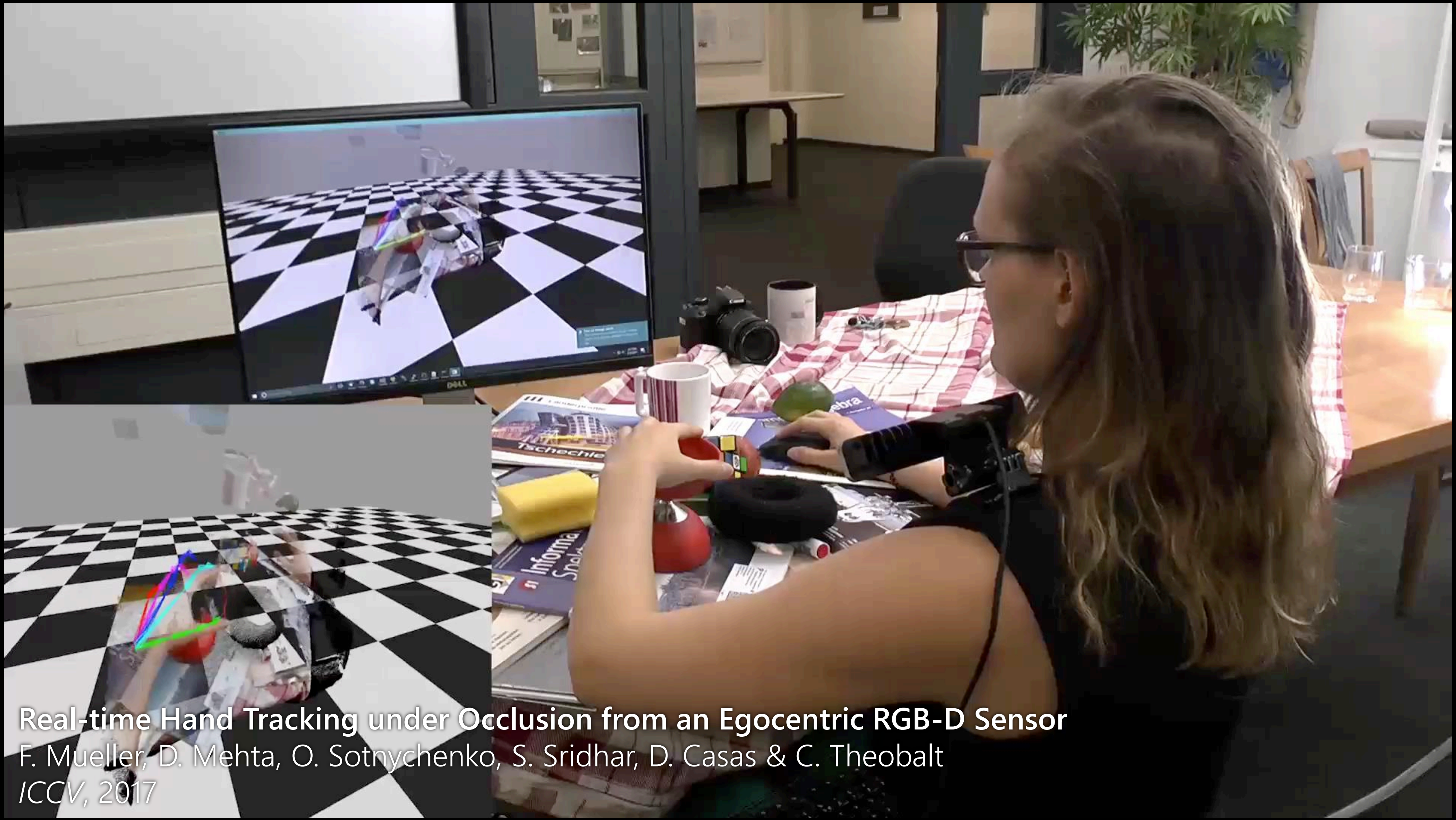
- **Goal:** reconstruct full hand pose (global transform + joint angles) using a single body-mounted camera
- Robust to:
  - fast and complex motions
  - background clutter
  - occlusions by arbitrary objects as well as the hand itself
  - self-similarities of hands
  - fairly uniform colour
- In real time (>30 Hz)



© F. Mueller et al.

Slide adapted from Fransizka Mueller

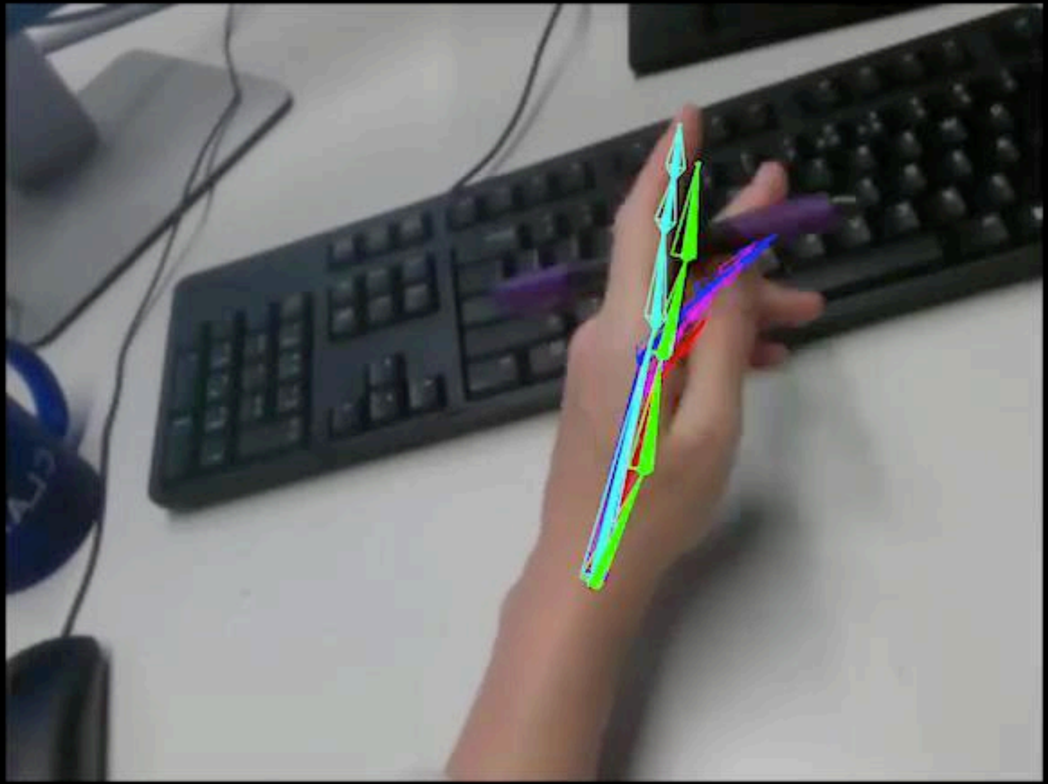
Egocentric hand tracking from RGB-D

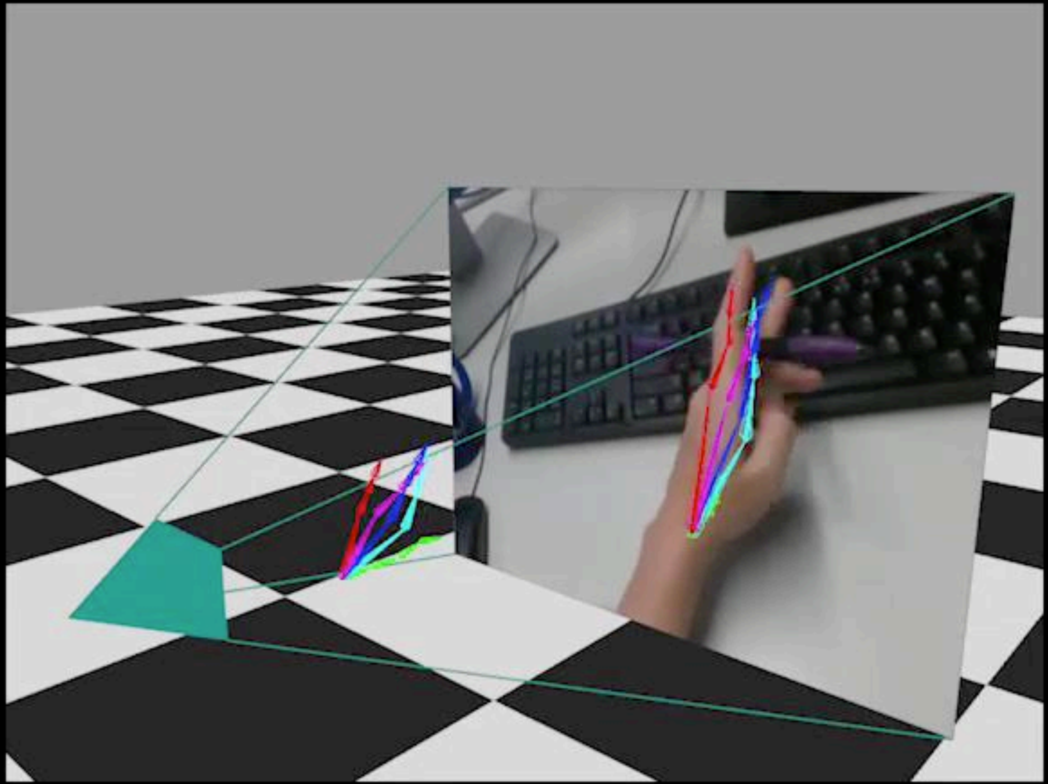


Real-time Hand Tracking under Occlusion from an Egocentric RGB-D Sensor
F. Mueller, D. Mehta, O. Sotnychenko, S. Sridhar, D. Casas & C. Theobalt
ICCV, 2017

https://youtu.be/Aay3MgHqQ0k?t=296

Egocentric hand tracking





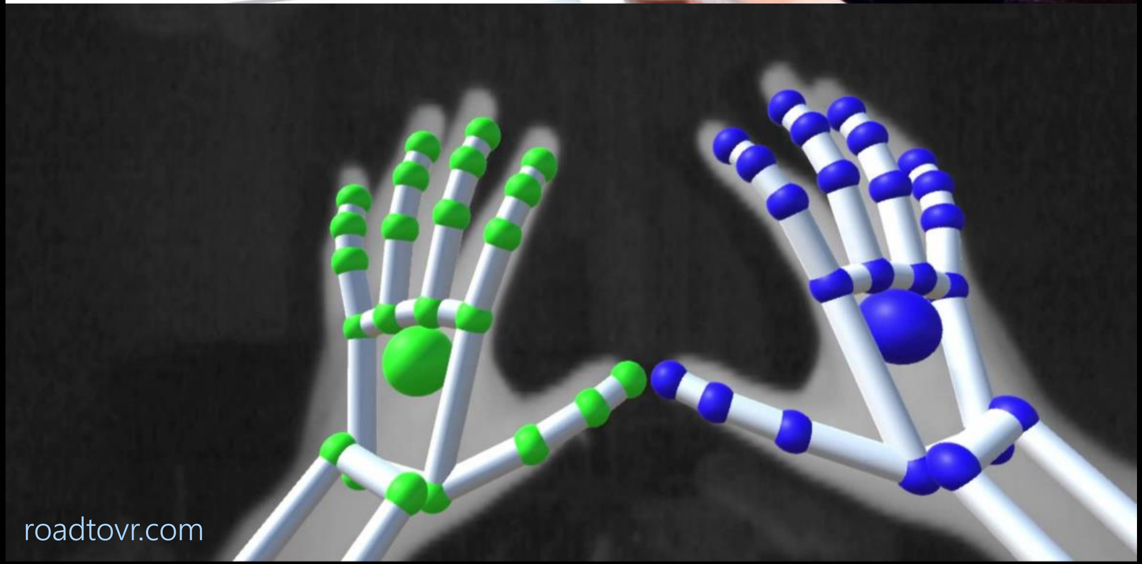
https://youtu.be/0wH0b9MdjPI?t=4

GANerated Hands for Real-time 3D Hand Tracking from Monocular RGB
F. Mueller, F. Bernard, O. Sotnychenko, D. Mehta, S. Sridhar, D. Casas & C. Theobalt
CVPR, 2018

Remaining challenges of hand tracking

- Robust results out of the box:
  - interacting with unknown objects
  - two hands simultaneously
  - no explicit model fitting
- Usability challenges:
  - not having sense of touch
  - line of sight required to sensor
  - fatigue from holding hands in front of sensor

NimbleVR

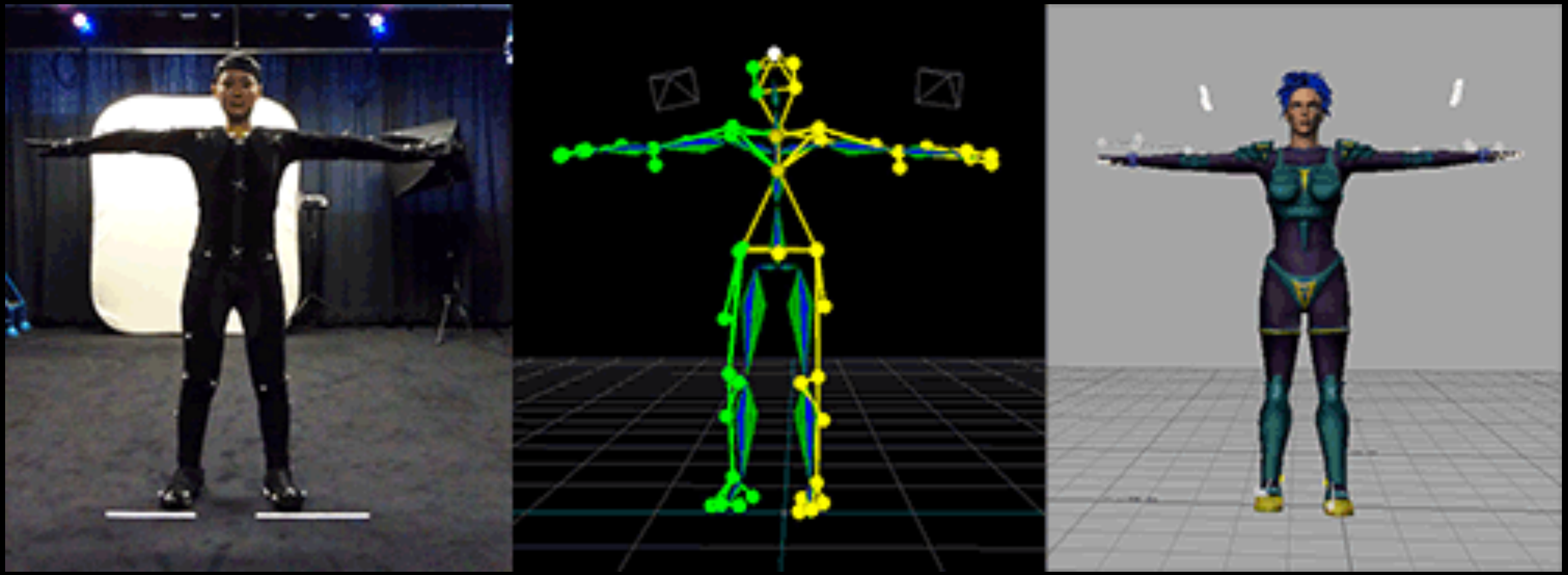


roadtovr.com

Slide adapted from Bruce Thomas & Mark Billinghurst

Full-body tracking

- Adding full-body input into VR:
  - creates illusion of self-embodiment
  - significantly enhances sense of presence



protchenko.info

Slide adapted from Bruce Thomas & Mark Billinghurst

Camera-based motion capture

- Use multiple cameras (8+) with infrared (IR) LEDs
- Retro-reflective markers on body clearly reflect IR light
- For example Vicon, OptiTrack:
  - – very accurate: <1 mm error
  - – very fast:
    - 100–360 Hz sampling rate
    - <10 ms latency
  - – each marker needs to be seen by at least two cameras

Vicon

OptiTrack

digitalcinema.com.ua

Slide adapted from Bruce Thomas & Mark Billinghurst

SIGGRAPH ASIA 2016 MACAO

EgoCap: Egocentric Marker-less Motion Capture with Two Fisheye Cameras

Helge Rhodin[1] Christian Richardt[123] Dan Casas[1],
Eldar Insafutdinov[1] Mohammad Shafiei[1]
Hans-Peter Seidel[1] Bernt Schiele[1] Christian Theobalt[1]

Graphics Vision & Video [1] mpi max planck institut informatik [2] intel VISUAL COMPUTING INSTITUTE [3] UNIVERSITY OF BATH

Today’s motion-capture challenges

- General environments
- Large scale motions
- Constrained rooms
- Easy to use, non-intrusive
- Low delay

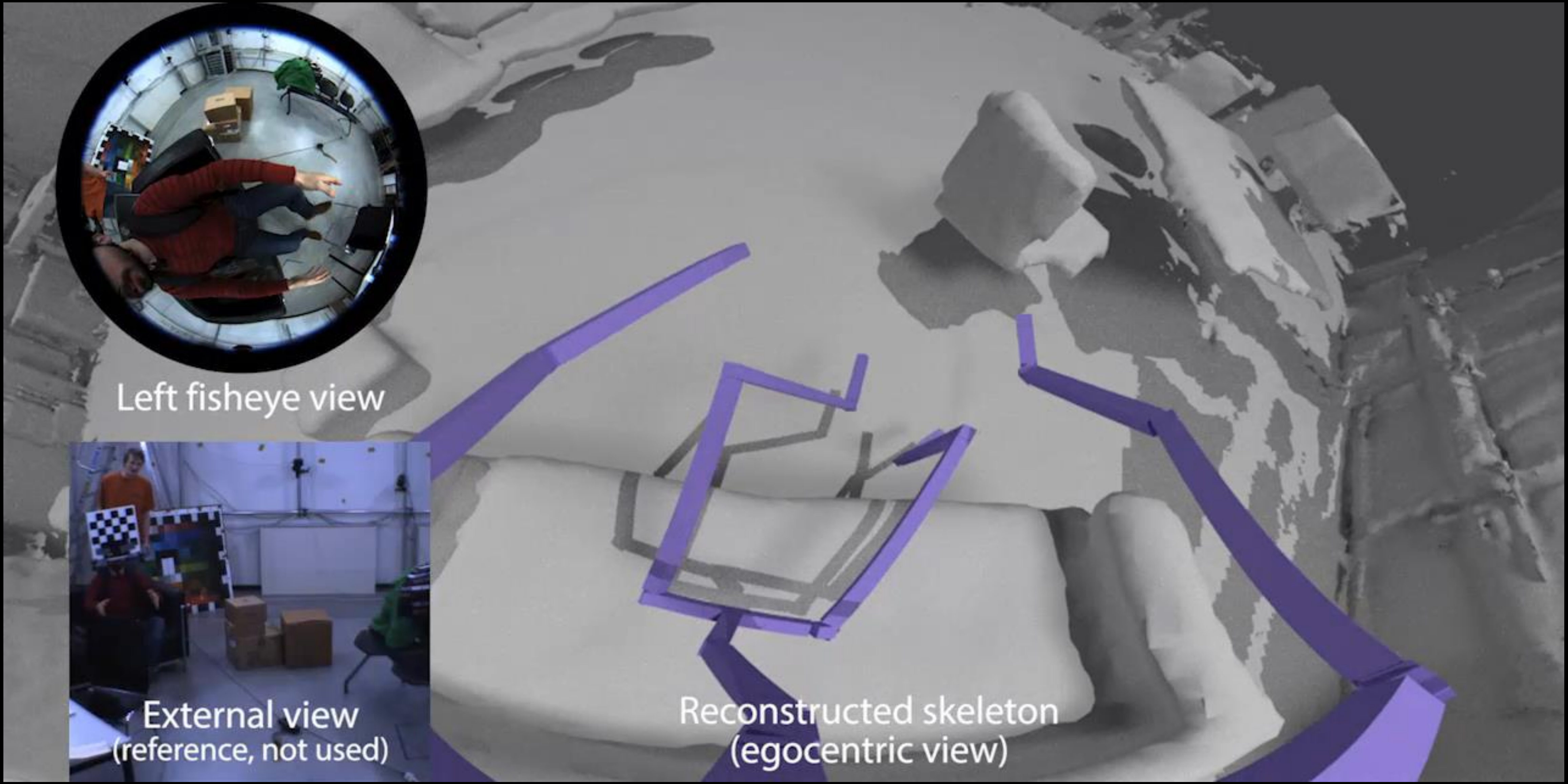
Computer animation

Sports and medicine

Autonomous driving

Virtual and augmented reality

Embodied virtual reality



Marker-less motion capture

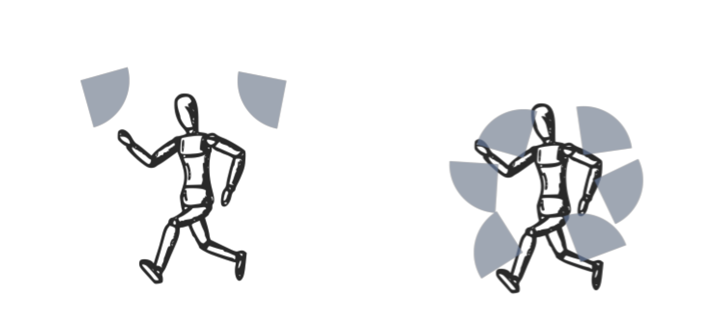
Outside-in

Non-intrusive

Limited capture volume

Full-body

Marker-less motion capture



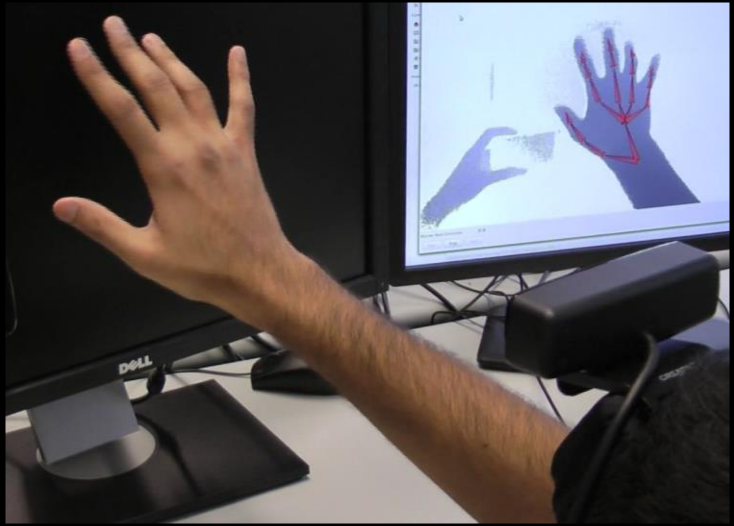
Outside-in	Inside-out
Non-intrusive	Intrusive
Limited capture volume	Infinite capture volume
Full-body	Full-body

[Shiratori 2011]

Marker-less motion capture

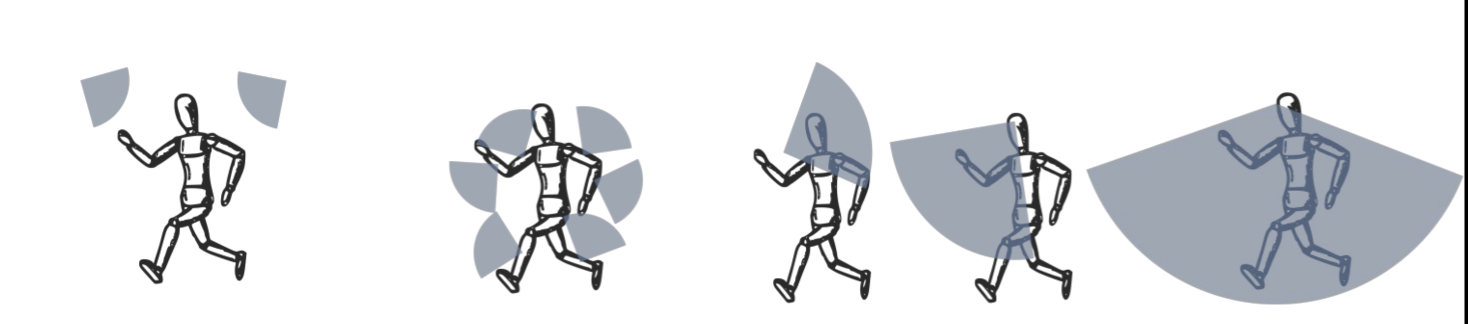
Outside-in	Inside-out	Inside-in
Non-intrusive	Intrusive	Low intrusion
Limited capture volume	Infinite capture volume	Infinite capture volume
Full-body	Full-body	Partial-body

[Jones 2011, Wang 2016]



[Sridhar 2015, ...]

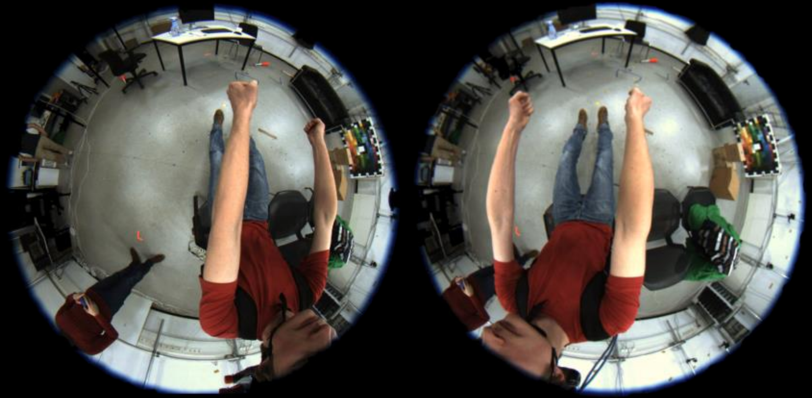
Marker-less motion capture



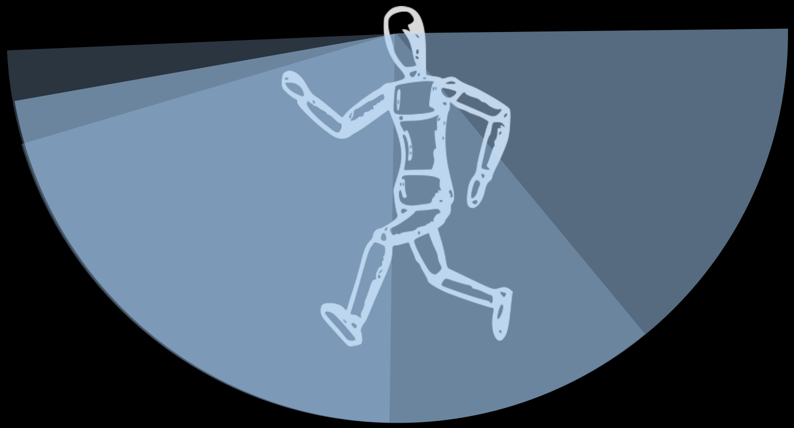
Outside-in	Inside-out	Inside-in	EgoCap
Non-intrusive	Intrusive	Low intrusion	Low intrusion
Limited capture volume	Infinite capture volume	Infinite capture volume	Infinite capture volume
Full-body	Full-body	Partial-body	Full-body

Camera gear

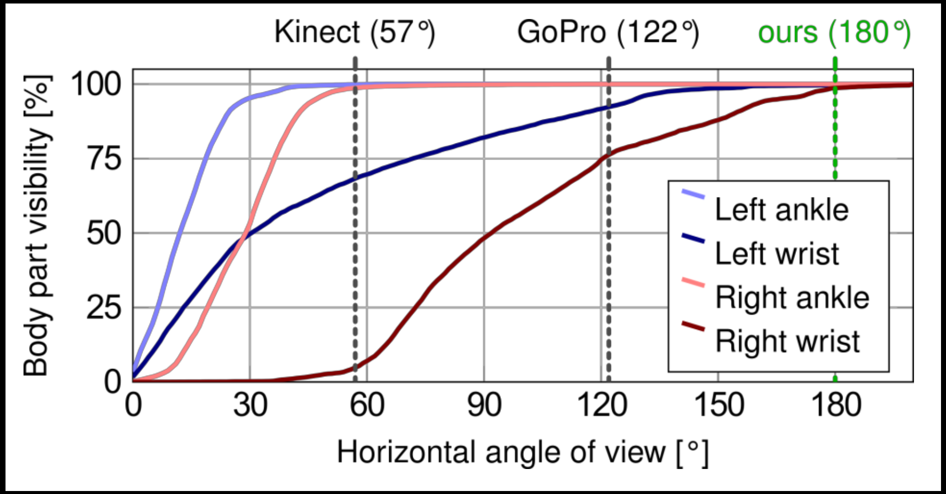
Camera extensions



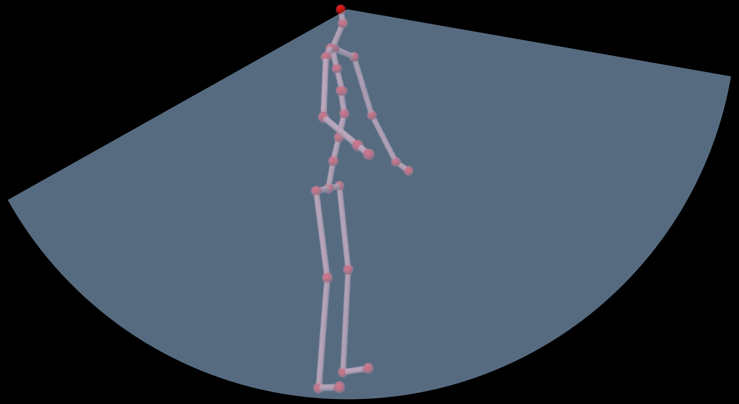
Egocentric view examples



Field of view



Egocentric capture challenges



Camera is attached

Subject is always in view

Human pose is independent of global motion

Estimation of global motion

Moving background

Top-down view

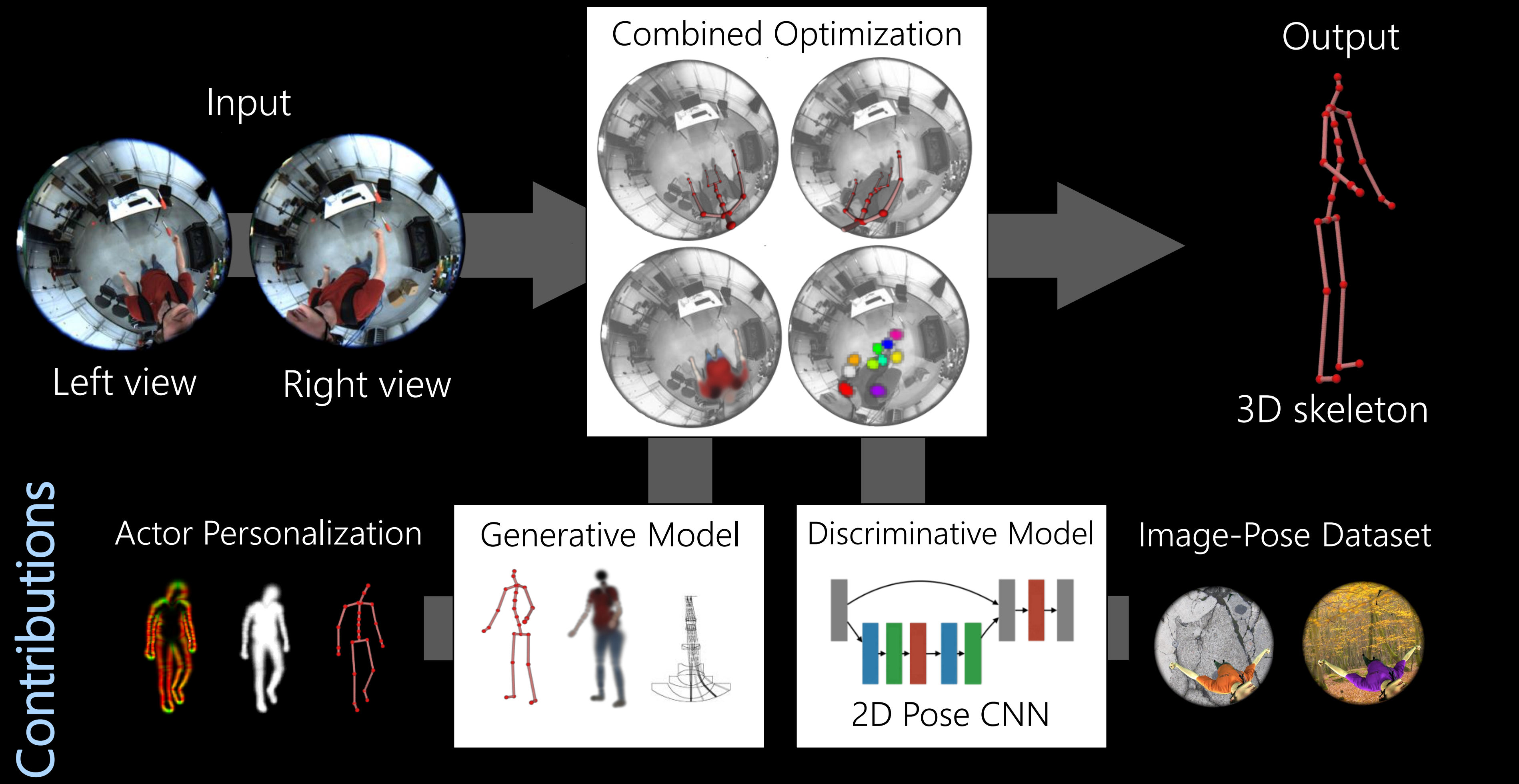
Self-occlusions

The lower body appears tiny

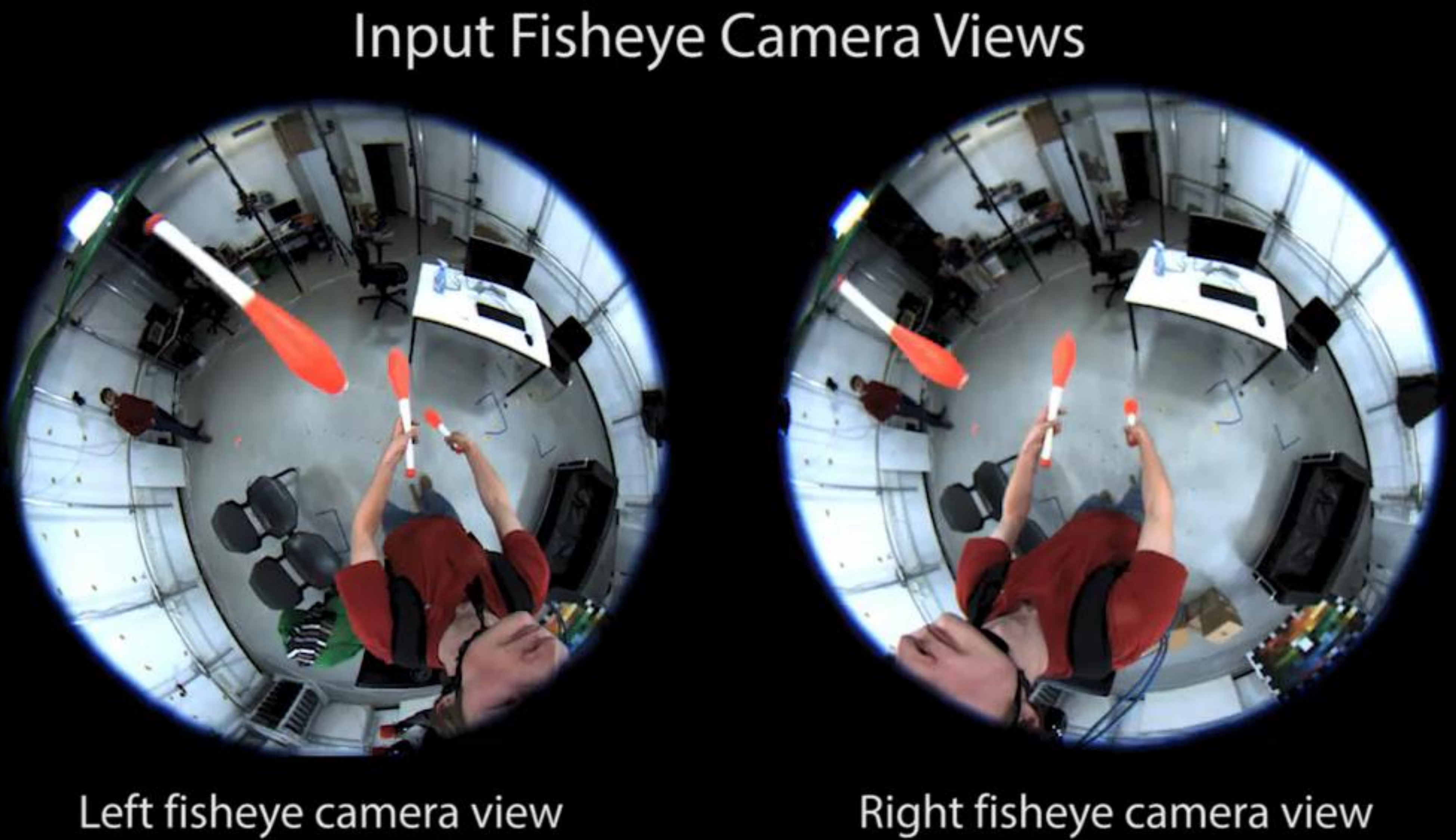
RGB only

Depth ambiguities

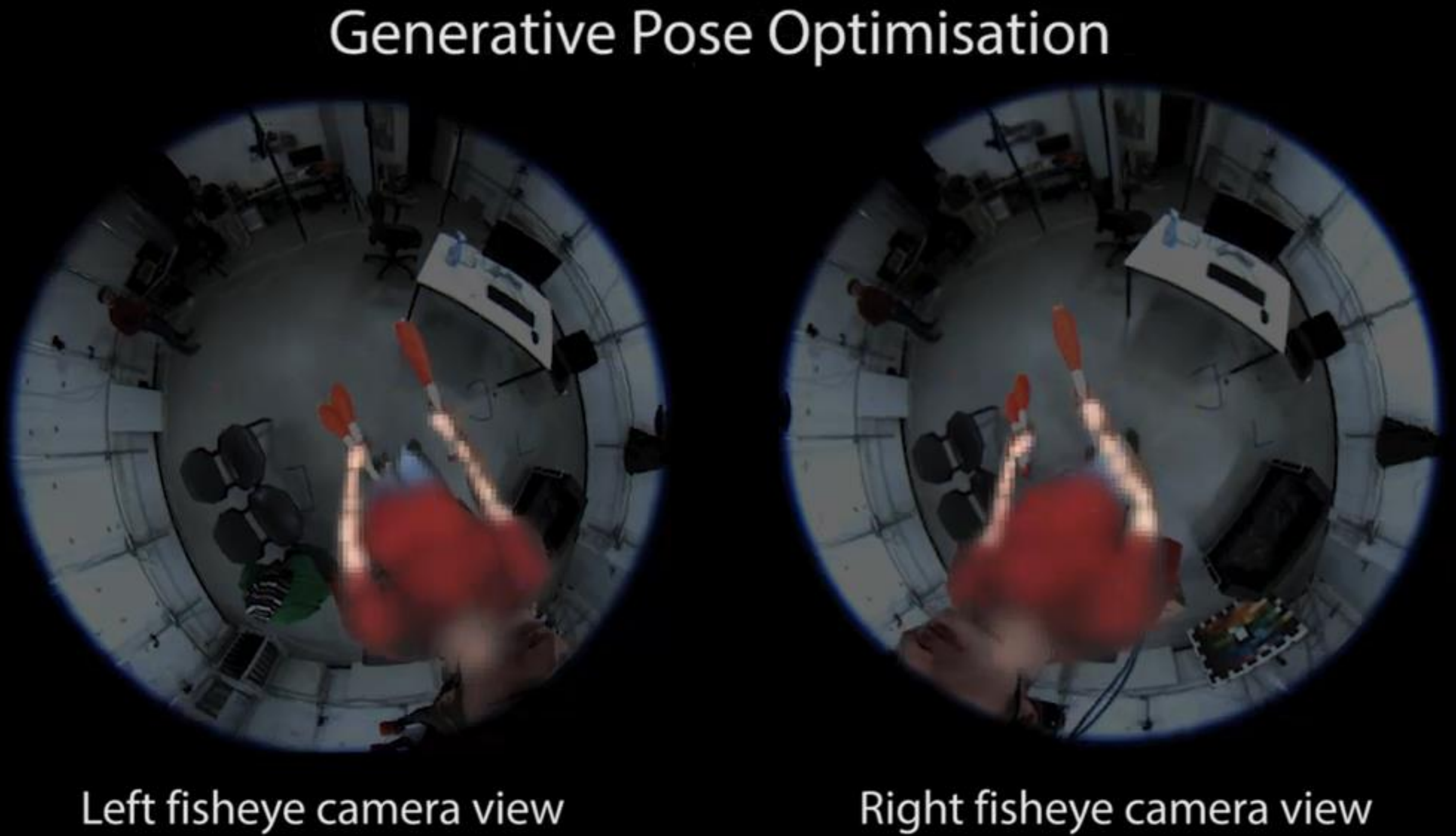
Model overview



Method walkthrough



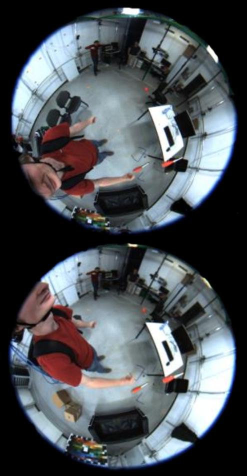
Method walkthrough



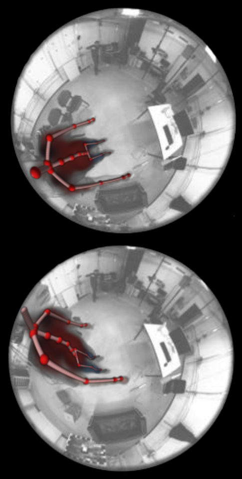
Combined optimization

- Energy minimization:
  - gradient descent on pose $\mathbf{p}^t$ at time t

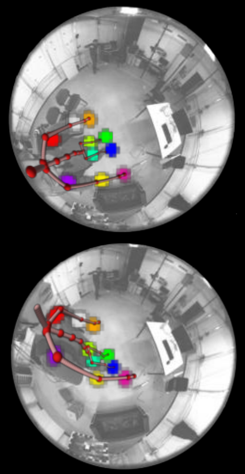
$$E(\mathbf{p}^t) = E_{\text{color}}(\mathbf{p}^t) + E_{\text{detection}}(\mathbf{p}^t) + E_{\text{pose}}(\mathbf{p}^t) + E_{\text{smooth}}(\mathbf{p}^t)$$



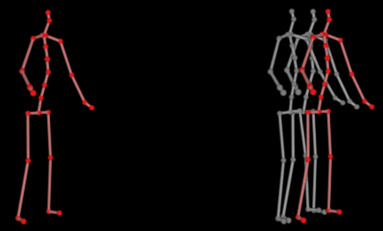
Input



Generative



Discriminative

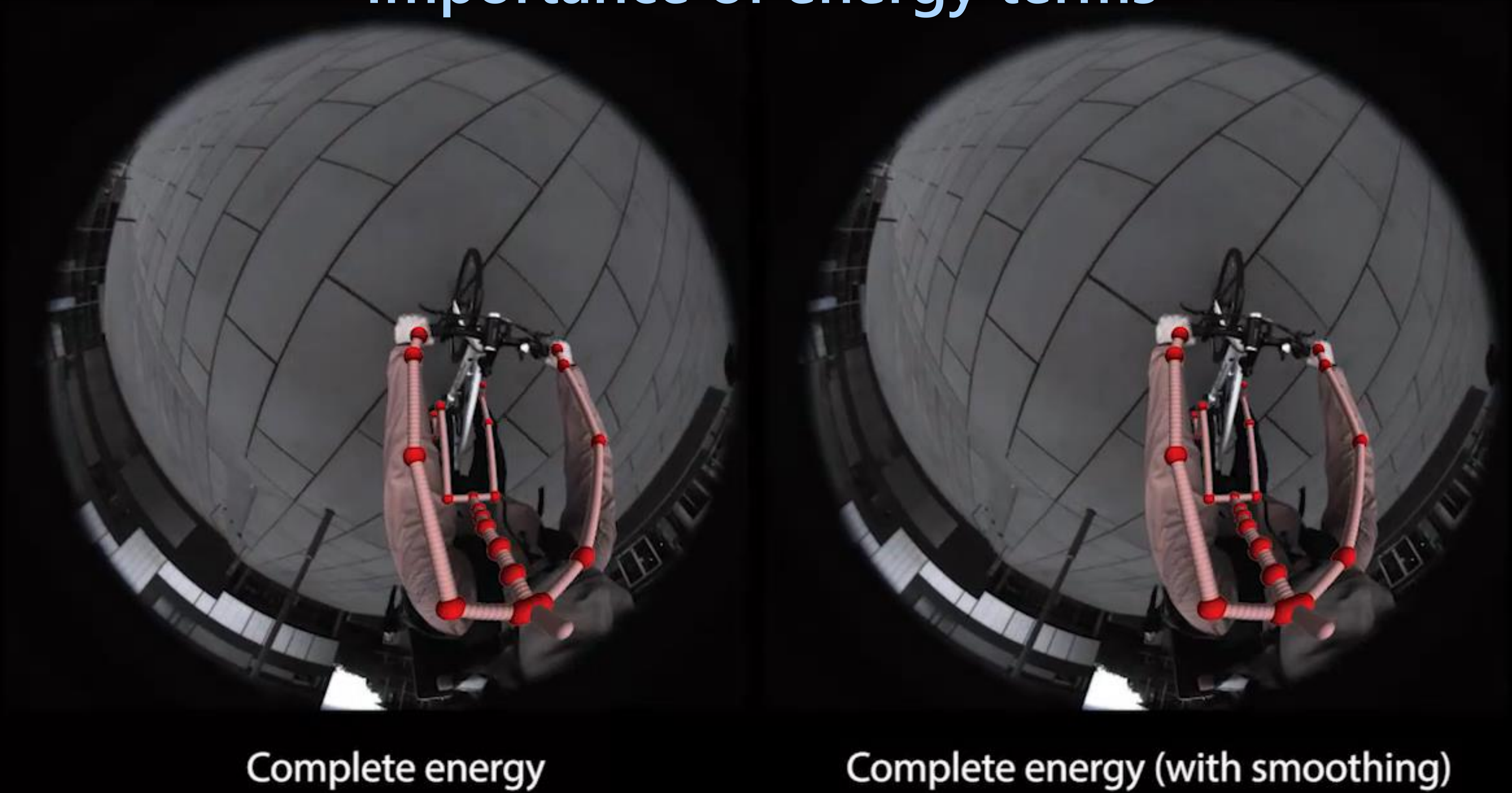


Prior terms

Importance of energy terms

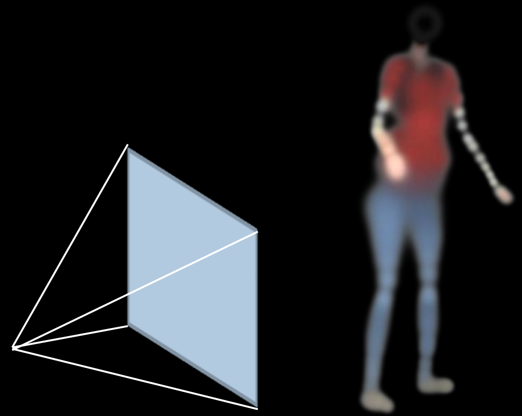
Without body-part detection term (Section 4.3.3)

Importance of energy terms

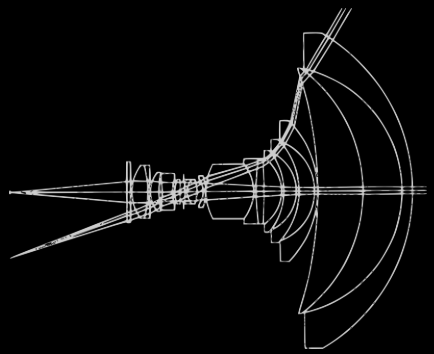


Generative model

- Volumetric body model
  - raytracing-based
  - fisheye camera
  - parallel GPU implementation



[Rhodin ICCV 2015, ECCV 2016]



[Scaramuzza 2006]

Our model

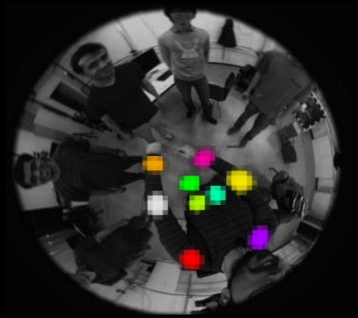
Discriminative component

- Deep 2D pose estimation
  - High accuracy with sufficient training data
  - Standard CNN architecture (Residual network [He 2016])

[Insafutdinov 2016, …]

- Egocentric training data?

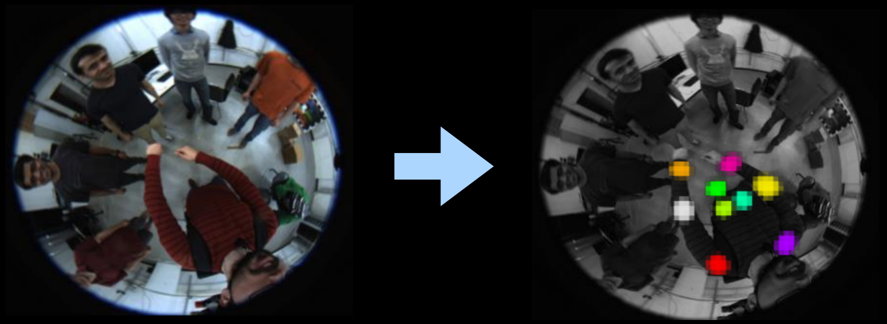
Example image



Annotation

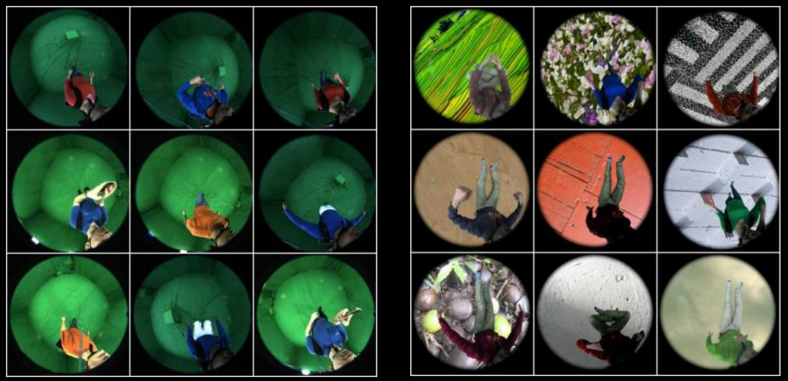
Training dataset

- Egocentric image-pose database
  - 80,000 images
  - appearance variation
  - background variation
  - actor variation



Example image

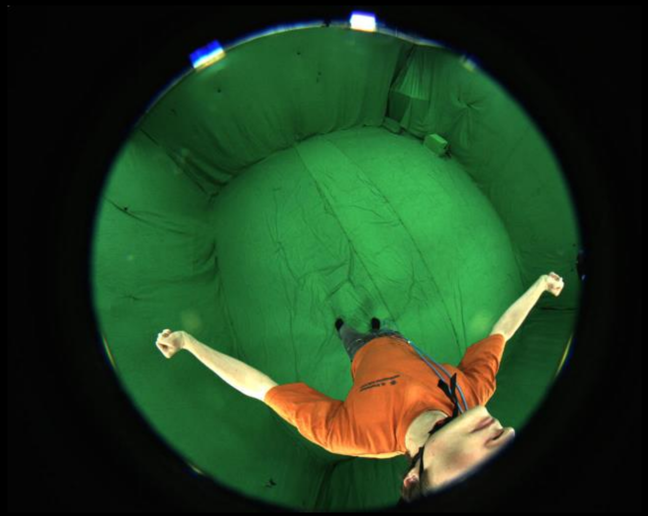
Annotation



Data augmentation

Ground-truth annotation

Diversity by augmentation: background

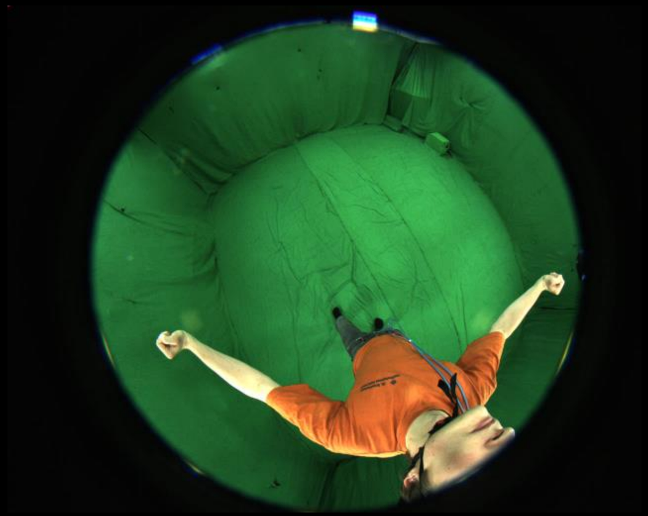


Original

Replaced background

- Green-screen keying to replace backgrounds
  - using random images from Flickr

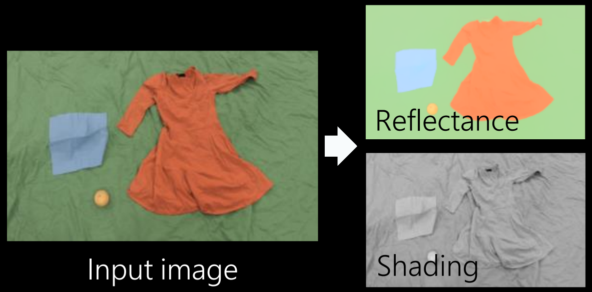
Diversity by augmentation: foreground



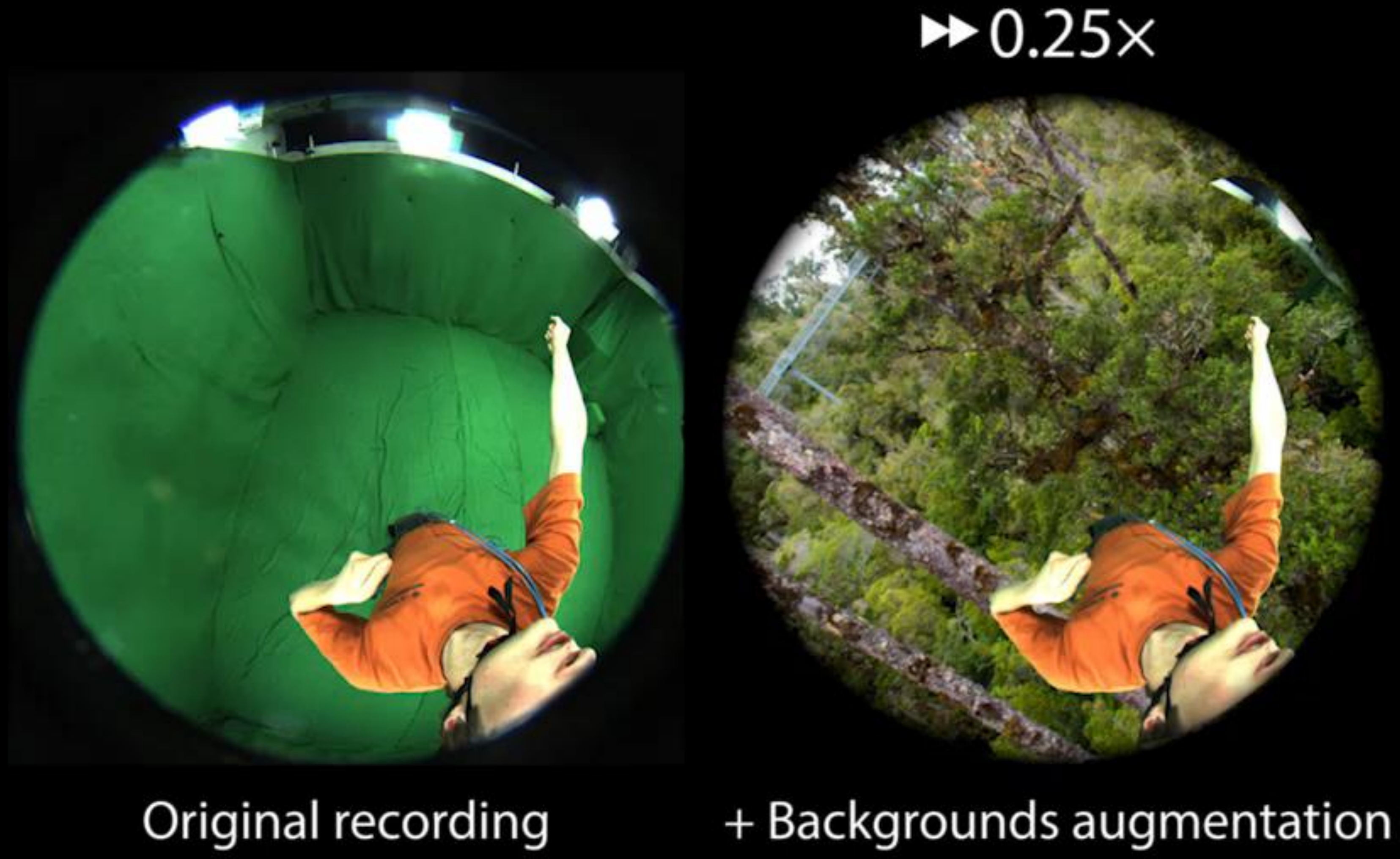
Original

Replaced albedo

- Intrinsic image decomposition [Meka 2016, …]

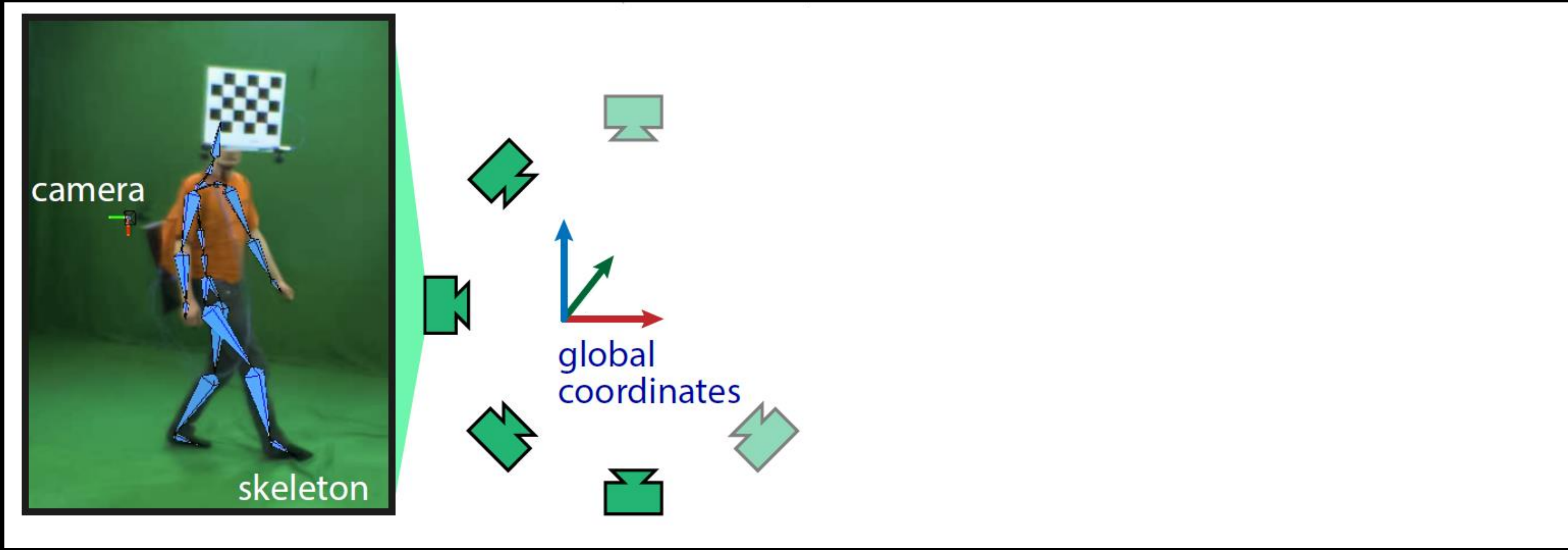


Training dataset augmentation



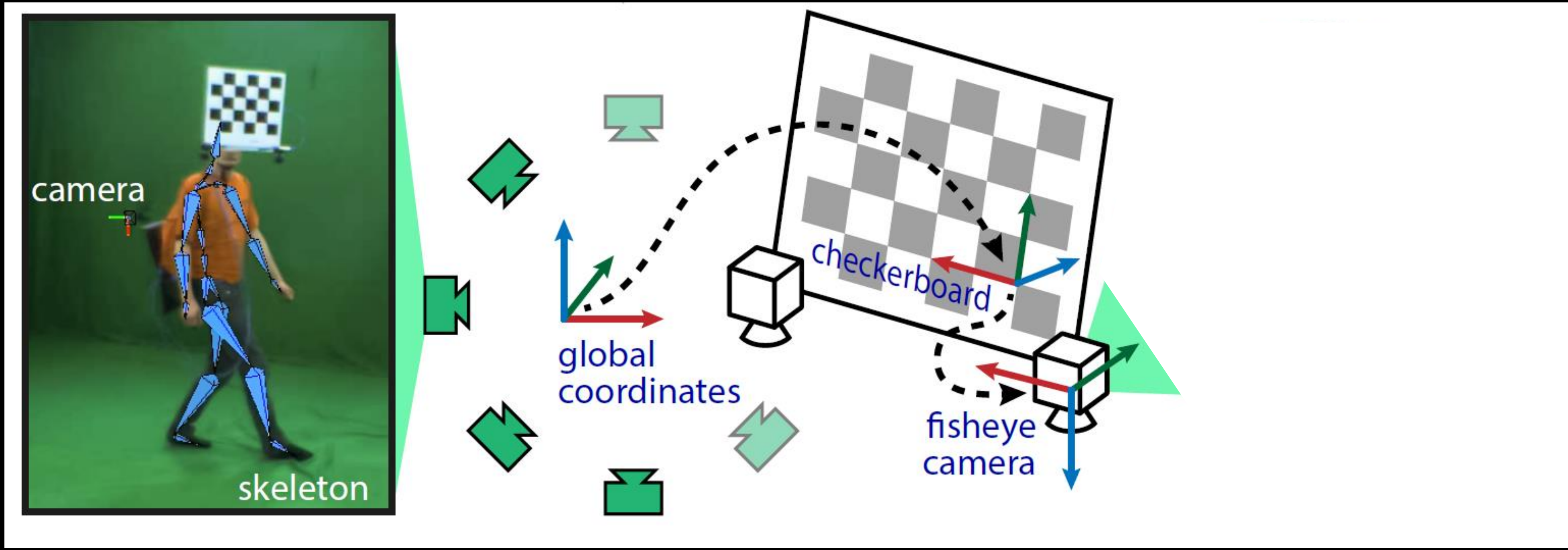
Automatic ground-truth annotation

Outside-in markerless motion capture



Automatic ground-truth annotation

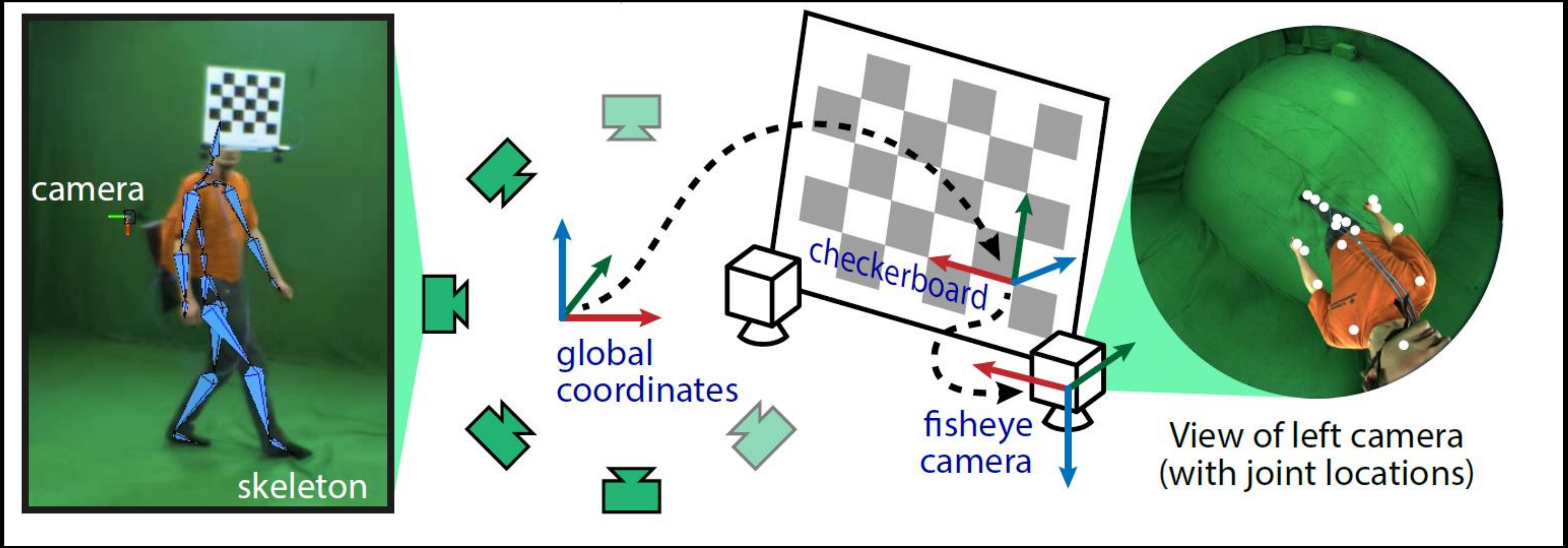
Outside-in markerless motion capture



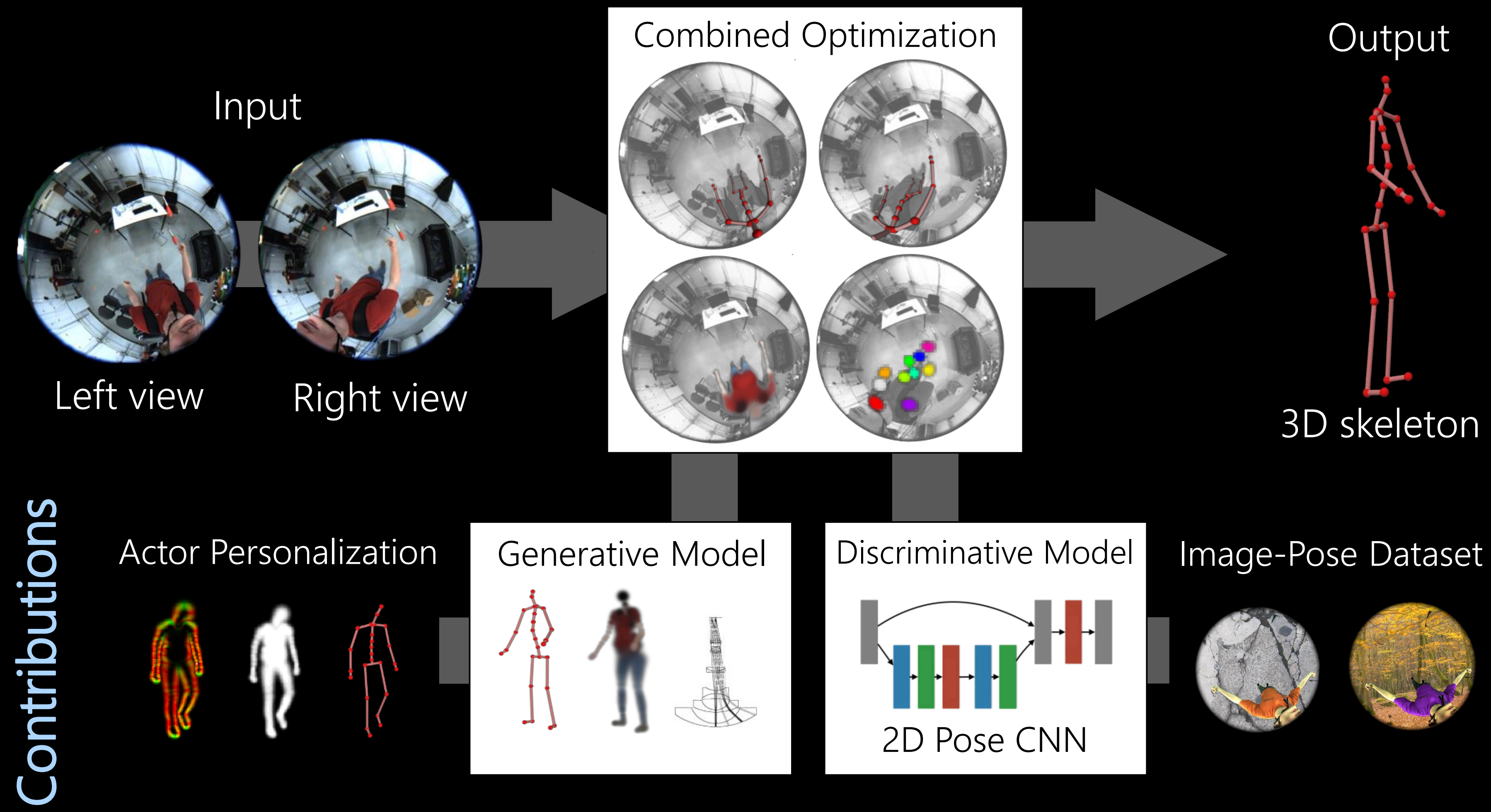
Automatic ground-truth annotation

Outside-in markerless motion capture

Projection into dynamic egocentric camera



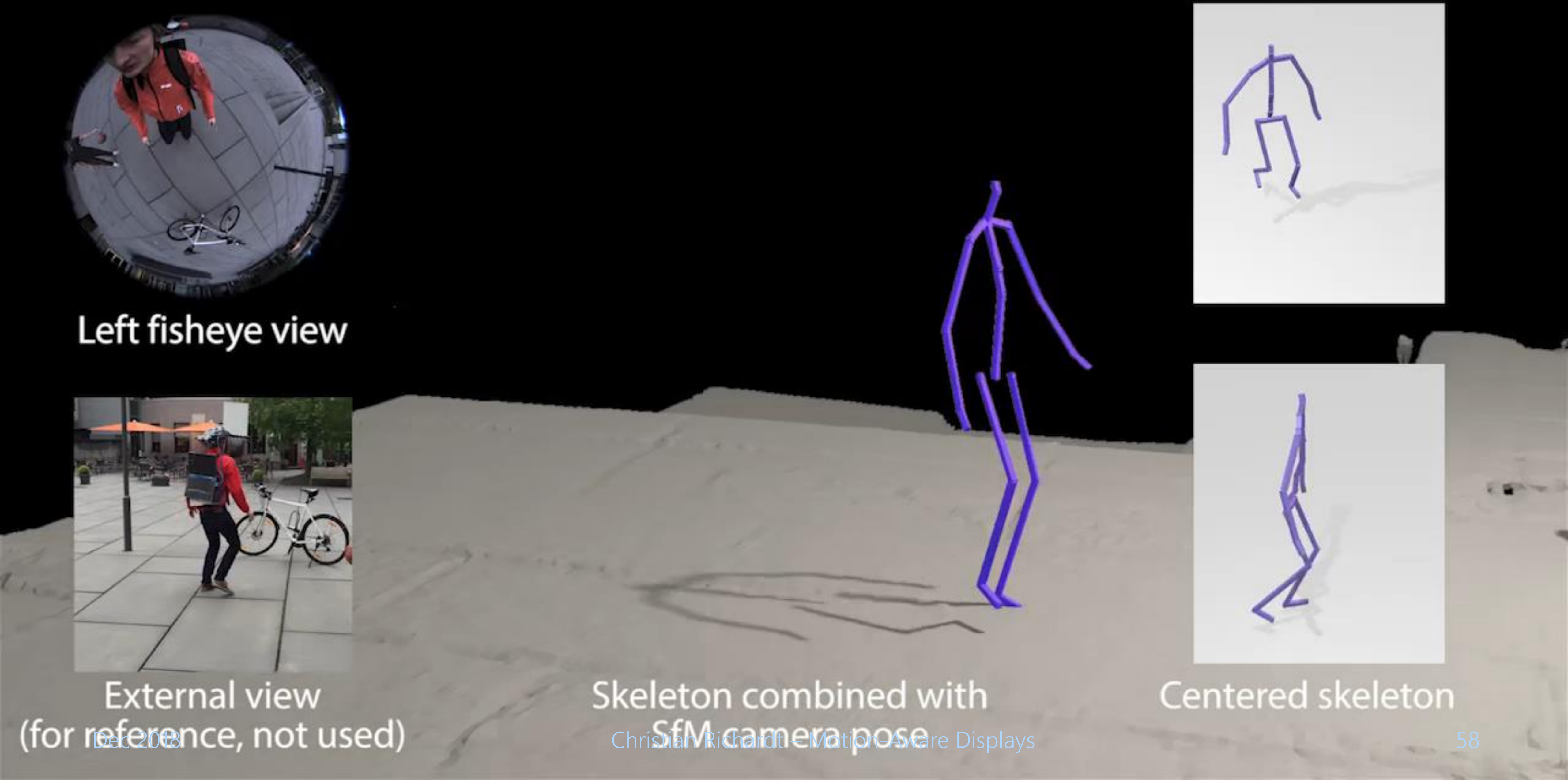
Model overview



Constrained and crowded Spaces


Two representative external views – Note the strong occlusions

Outdoor and large-scale

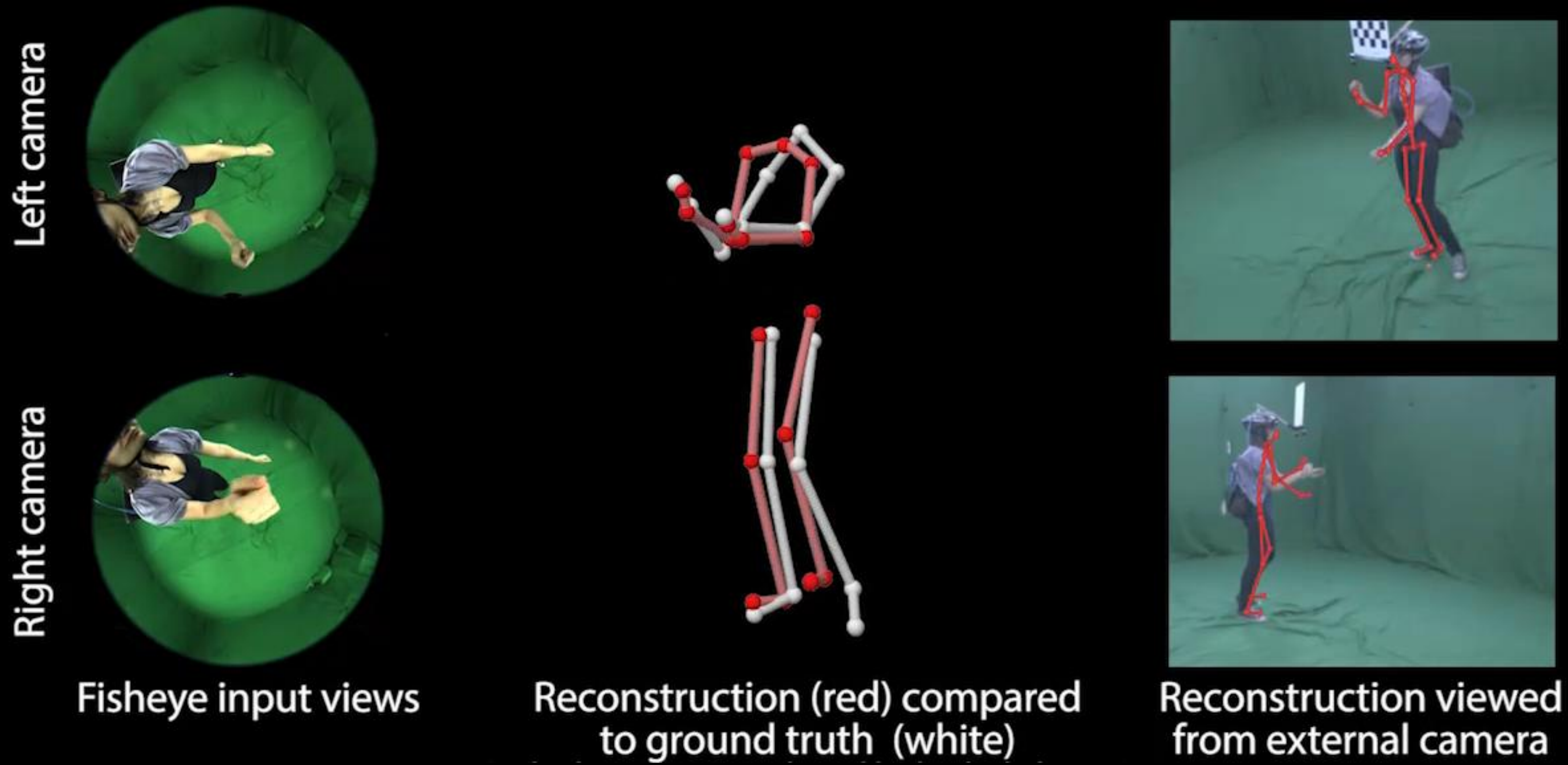


Virtual and augmented reality

Embodied virtual reality

Quantitative analysis

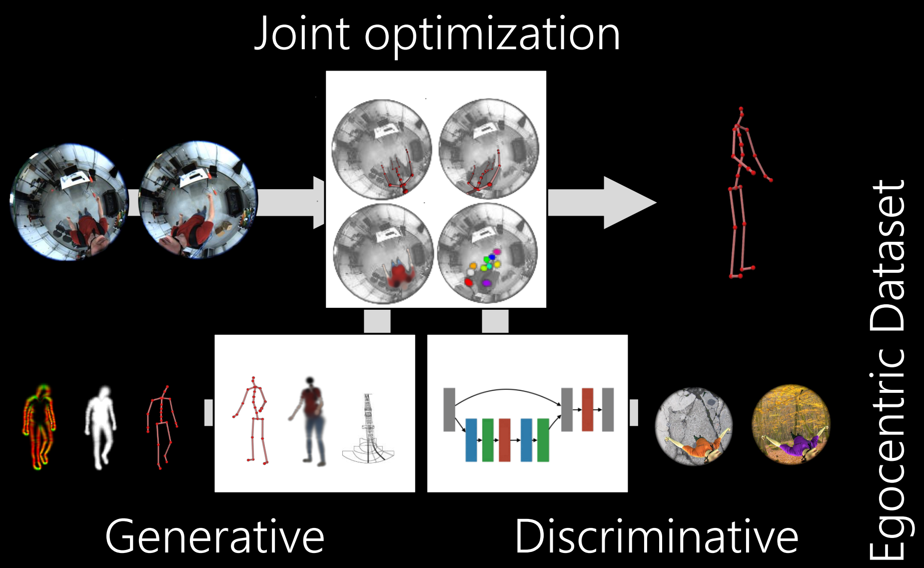
- 7 cm average Euclidean 3D error
- Temporally stable



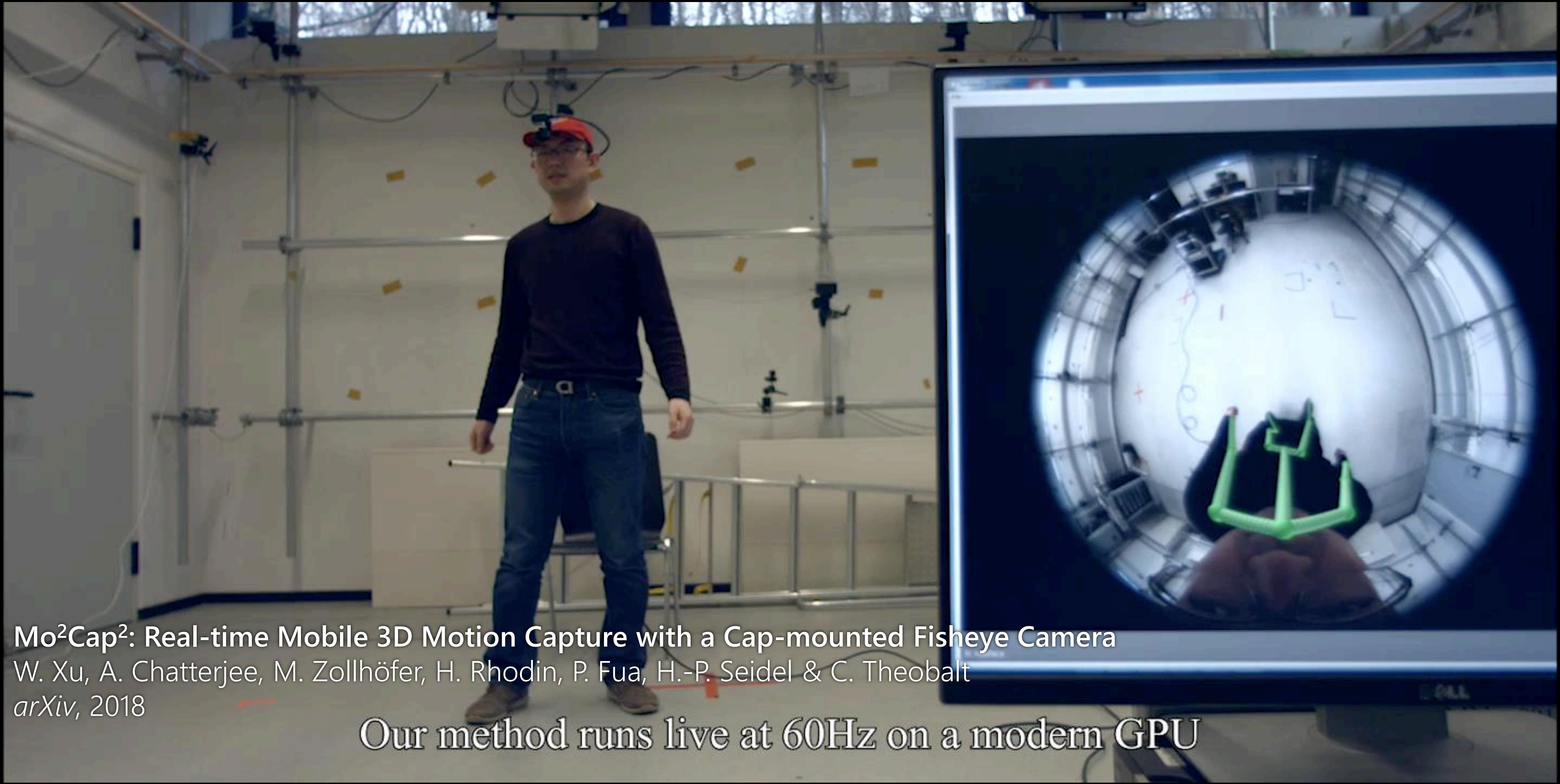
Occlusions – limitations

EgoCap summary

- Inside-in motion capture
  - full-body 3D pose
  - easy-to-setup
  - low intrusion level
  - real-time capable
  - general environments
- Future work
  - low latency (for VR)
  - alternative camera placement, monocular
  - capture hands and face



Single-camera egocentric motion capture



Mo²Cap²: Real-time Mobile 3D Motion Capture with a Cap-mounted Fisheye Camera
W. Xu, A. Chatterjee, M. Zollhöfer, H. Rhodin, P. Fua, H.-P. Seidel & C. Theobalt
arXiv, 2018

http://gvv.mpi-inf.mpg.de/projects/wxu/Mo2Cap2/

Quick recap

- Immersion & presence: motion is extremely important
  - presence breaks when visual body motion does not match physical motion
- Tracking in VR/AR: need high accuracy and update rate, low latency
  - in practice, usually best to combine IMUs with optical tracking to fix drift
- Hand input devices: controllers are tracked robustly and accurately
  - hand tracking will soon enable natural interaction with real-world objects
- Full-body motion capture: bring the entire body into VR
  - marker-based systems are fast, robust, accurate and very expensive
  - markerless systems allow live motion capture from just 1 or 2 cameras

Questions?

Christian Richardt

Motion-Aware Displays

SIGGRAPH Asia Course on Cutting-Edge VR/AR Display Technologies

CAMERA
Centre for the Analysis of Motion, Entertainment Research and Applications

UNIVERSITY OF BATH

richardt.name
c_richardt